College

PHYSICS

FOR

COLLEGE STUDENTS

BY

HENRY S. CARHART, Sc.D., LL.D.
PROFESSOR EMERITUS OF PHYSICS, UNIVERSITY OF MICHIGAN

ALLYN AND BACON
Boston and Chicago

~~Physics Dept~~

FDP

Norwood Press
J. S. Cushing Co. — Berwick & Smith Co.
Norwood, Mass., U.S.A.

PREFACE

This book is a response to many inquiries for a text somewhat more advanced than the *High School Physics* of this series and distinctly less mathematical than the *University Physics*. In writing it the author has kept in mind those students who are not necessarily scientific in their taste or choice, but who desire a comprehensive outline of the essential principles of Physics. The value of this subject is too generally acknowledged to admit of question, but the supposed mathematical difficulties are often assumed to be a formidable barrier to its pursuit. The author confidently believes that he has been successful in his effort to reduce the difficulties to such a degree that they may readily be surmounted by the average college student. This is quite a different problem from that of making a book elementary by eliminating all the more difficult parts, whether they are essential or not. It must be left to the author's fellow teachers to decide whether he has succeeded in securing simplicity without the sacrifice of quality.

September, 1910.

CONTENTS

MECHANICS

SOUND

LIGHT

HEAT

Chapter XIII. Nature and Effects of Heat

Chapter XIV. Transmission and Radiation of Heat

Chapter XV. Thermodynamics

MAGNETISM AND ELECTRICITY

Chapter XVI. Magnets and Magnetic Fields

Chapter XVII. Electrostatics

Chapter XVIII. Electric Currents

COLLEGE PHYSICS

MECHANICS

CHAPTER I

INTRODUCTORY

I. MATTER AND ENERGY

1. Reality of Physical Phenomena. — In studying the facts of nature we assume the objective reality of physical phenomena and the existence of external objects apart from the mind of the observer. While we become acquainted with the physical universe solely through our senses, it will probably be admitted that every form of matter, such as a pebble, a drop of dew, and the oxygen of the air, has objective existence. The supreme test of physical reality is the fact that the material world remains unchanged in quantity in whatever way it is measured. From this point of view, only two classes of things or entities are found in the physical world, — *matter* and *energy*.

2. Matter. — The ultimate nature and structure of matter are not known with any degree of certainty. The discovery of radium, and its spontaneous disintegration into substances of simpler constitution, such as helium, and of the generation of heat by its disintegration, have profoundly modified the older views of the constitution of matter. Matter is often defined as something occupying space, but it is much better described by its properties.

A limited portion of matter is a *body*, and different kinds of matter, having distinct properties, are called *substances*.

A copper coin, a drop of rain, the air in a pneumatic tire, are bodies. Copper, water, and air are substances, since each has properties distinguishing it from all others.

3. Energy. — There is every reason to believe that the motion of a body cannot be altered unless motion is imparted to it by another body or system of bodies by some such method, for example, as collision. It is in accordance with the usual mode of expression in such cases to say that the second body *does work* on the first. *Energy* may be defined as *the capacity for doing work*, where work may be considered as the act of changing the motion, the relative position, or the chemical and physical constitution of another body against a resistance opposing the change. Thus, work is done in winding a clock by coiling a spring against the reaction of bending, or by lifting a weight against the resistance of gravity. The coiled spring or the lifted weight then possesses energy, and it may in turn do work by giving motion to the pendulum and keeping it swinging against friction and the resistance of the air.

Energy, like matter, has the property of *conservation*, and must be regarded as having real existence. It may be transformed and passed on in an endless series of changes, but it remains unchanged in amount.

4. Definition of Physics. — Physics is the science which treats of the related phenomena of matter and energy. It must not be assumed that physics is sharply differentiated from the other physical sciences by well-defined boundaries. A branch of science is classified in accordance with its chief aims. Chemistry, for example, endeavors to ascertain how the so-called elements enter into the composition of bodies, and the laws governing those changes in matter which affect its properties and identity. The combustion of carbon, the rusting of iron, the conversion of limestone into lime by heat, the fermentation of wine, are all changes in the constitution of the individual molecules of the several bodies.

They are therefore chemical changes. At the same time all of them involve changes in the associated energy, and this fact gives them a physical character. The old distinctions between physics and chemistry are highly artificial, and the discovery of new phenomena has the effect of making those distinctions obsolete. In fact, so obscure is now the border line between them that a new subdivision, called physical chemistry, which lies partly in the one field and partly in the other, is recognized as a distinct branch of science.

II. PROPERTIES OF MATTER

5. General and Special Properties. — The properties of matter are the qualities which serve to describe it and to define it provisionally. These properties are either *general*, that is, common to all kinds of matter in whatever state it may exist; or *special*, those distinctive of some kinds of matter and conspicuously absent in others. A property shared by all kinds of matter alike is *extension*, which means that every body occupies space or has dimensions. Another general property is *impenetrability*, which means that two bodies cannot occupy the same space at the same time.

On the other hand, a piece of clear glass lets light pass through it, or is *transparent;* while a piece of sheet iron does not transmit light, or is *opaque*. A watch spring recovers its shape after bending and is *elastic;* while a strip of lead possesses this property in so slight a degree that it is classed as *inelastic*. Extension and impenetrability are general properties of matter, while transparency and elasticity are special properties.

6. Inertia. — The most general and characteristic property exhibited by all matter is *inertia*. In fact, inertia is the only property inherent in matter which has to do with motion. Inertia is the persistence of matter in its state of either rest or motion, and its resistance to any attempt to

change that state. If a moving body stops, its arrest is always due to some influence outside of itself; and if a body at rest be set in motion, the motion must be imparted to it by some other body.

Many familiar phenomena are due to inertia. When a fireman shovels coal into a furnace, he suddenly arrests the motion of the shovel and leaves the coal to move forward into the furnace by its inertia. Suspend a heavy body and strike it a blow with a light hammer; the hammer rebounds as if the heavy body were fixed in space by its inertia. A smooth cloth may be snatched from under a dish almost without disturbing it. The persistence with which a spinning top maintains the direction of its axis of rotation is due to its inertia. The violent jar to a water pipe when a faucet is suddenly closed is accounted for by the inertia of the stream. Tall columns and the detached portions of chimneys are sometimes twisted around by sudden gyratory earthquake movements (Fig. 1). The sudden circular motion of the earth under the column leaves it behind, and the slower return motion carries the column with it.

Fig. 1

Suspend a heavy weight by a string, as in Figure 2, using the same kind of string above and below the weight. A *steady* pull on the string at B will break the upper string because it carries the greater load. A *sudden* downward pull on B, however, will cause the lower string to break. On account of the inertia of the weight, the lower string breaks before the pull reaches the upper one.

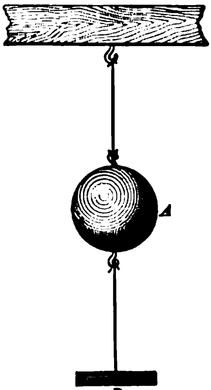


Fig. 2

7. **Mass.** — The *mass* of a body is the measure of its inertia, that is, of the resistance which a body offers to motion or change of motion. This resistance to motion must be

carefully distinguished from friction, resistance of the air, or any other opposition to motion or change of motion except the inherent resistance of inertia. Mass must not be confused with weight (§ 55), because mass is not dependent on gravity. The mass of a meteoric body is the same when flying through space as when it strikes the earth and embeds itself in the ground. If it could reach the center of the earth, its weight would be zero; at the surface of the sun, it would weigh nearly twenty-eight times as much as at the earth's surface; but its mass would be the same everywhere.

8. **Porosity.** — Sandstone, hardened plaster of Paris, unglazed porcelain, and many other bodies absorb a considerable quantity of water without appreciable change in volume. The water fills the interspaces which are visible either to the naked eye or under a microscope. These interspaces, or pores, whether visible or invisible, are not a part of the space occupied by the body, and therefore interpenetration is possible. All matter is probably porous in structure, and the corresponding property is called *porosity*.

We may regard all bodies as built up of a number of very small masses called *molecules*. In gases these molecules are separated by relatively large spaces, while even in solids and liquids the molecules are not in actual contact. Hence the possibility of interpenetration and compression. If, for example, 50 cm.^{3} of alcohol be mixed with 50 cm.^{3} of water, the volume of the mixture will not be 100 cm.^{3}, but only about 97 cm.^{3}. Interpenetration takes place to a certain extent. In a famous experiment of the Florentine Academicians, a hollow sphere of heavily gilded silver was filled with water and subjected to pressure. The water exuded through the pores of the silver and gold and stood in beads on the surface. Francis Bacon observed the same phenomenon with a lead sphere.

9. **Elasticity.** — Stretch a rubber band, bend a bow or a knitting needle, compress a tennis ball, twist a steel wire. In each case the form or volume has been changed, or both, and the body has been *strained*. A *strain* is either a change of shape or a change of volume due to a distorting force,

called a *stress*. The property of recovery from a strain when the stress is removed is called *elasticity*. When a body recovers its form after release from a force of distortion, it has *elasticity of form;* it has *elasticity of volume* when the temporary distortion is one of volume.

10. **Plasticity.** — The inability of a body to recover from distortion produced by a stress is called *plasticity*. In so far as bodies are not elastic they are plastic, and even elastic bodies are plastic beyond the limits where they cease to be perfectly elastic. Plastic bodies require force to change their shape, but the continued application of a stress to maintain the change is not necessary. Bodies are classed as elastic if they have a large limit of elasticity, and plastic if their limit of elasticity is small. A bar of lead will vibrate to a certain extent if struck, showing that it is somewhat elastic; but its elastic limit is so small that it is classed as a plastic body.

11. **Cohesion.** — The particles or molecules of a body are held together by *cohesion*. Cohesion unites the molecules together throughout the mass, whether the molecules be like or unlike. *Adhesion* unites bodies by their adjacent surfaces. Cohesion not only holds together the molecules to form a visible mass, but it resists a force tending to break or crush a body. Welding consists in bringing clean metallic surfaces into intimate molecular contact by heating and hammering so that they cohere. When a clean glass rod is dipped into water and then withdrawn, a drop adheres to its surface. Two freshly cut surfaces of lead will adhere if brought into close contact by pressure.

12. **Tenacity.** — *Tenacity* is the resistance which a body offers to being torn asunder. It is determined by finding the weight necessary to break it in the form of a round wire.

Tenacity diminishes with the duration of the pull. A smaller force applied for a long time will often break a wire which would not be broken by a larger force of short duration.

Lead has the least tenacity of all solid metals, and cast steel the greatest; yet the latter is exceeded by fibers of silk and cotton. Single fibers of cotton can support millions of times their own weight. The tenacity of wood is greater along the fibers than across them.

13. Ductility and Malleability. — *Ductility* is the property of a substance which permits it to be drawn into wires or filaments. Platinum is the most ductile of all metals. Wollaston drew a wire of it only 0.00003 of an inch in diameter. A mile of this wire would have weighed only 1.25 grains.

Some substances become highly ductile only at high temperatures. Molten glass has been spun into a thread less than 0.0001 of an inch in diameter; a mile of it would weigh only a third of a grain. Professor Boys has drawn fibers from white-hot quartz not more than 0.00001 inch in diameter. Such quartz threads have a tenacity approaching that of steel.

Malleability is a modification of ductility which permits some metals to be hammered or rolled into thin sheets.

Pure gold is more malleable than any other substance. It has been hammered between skins to a thickness of only 0.00003 inch. Other metals possess this property, but to a less degree. Zinc is malleable when heated to a temperature of from 100° to 150° C. and it can then be drawn into wire or rolled into sheets. At 210° C. it again becomes brittle and can be crushed to powder. Nickel at red heat can be worked like wrought iron.

Articles cast from pig iron may be made somewhat malleable by heating for several days in contact with an agent, such as hematite, which removes some of the carbon from the cast iron.

14. Hardness. — *Hardness* is the resistance offered by a body to scratching or abrasion. It is a relative term only. Diamond is the hardest of all bodies because it scratches all others and is not scratched by any. It can be ground only by its own powder. Diamonds subjected to great hydraulic pressure between mild steel plates completely embed themselves in the metal. Hard bodies, such as emery and carborundum, are used as polishing powders and for grinding metals.

Steel becomes very hard and brittle when suddenly cooled from a high temperature. The process of *tempering* consists in the subsequent gradual reheating till the hardness is diminished to the required extent, as indicated by the color of a polished surface. Cutting instruments are made of tempered steel. An alloy of four parts copper and one part tin is ductile and malleable when rapidly cooled, but hard and brittle when cooled slowly.

15. States of Matter. — Matter exists in three distinct states of molecular aggregation, — the solid, the liquid, and the gaseous. All three are exemplified by water, which may be in the condition of ice, water, or water vapor. The difference is not one dependent entirely on temperature. All three forms may exist in contact at the same temperature as ice, ice-cold water, and vapor of water.

A *solid* has independent form and volume, and resists any stress tending to alter its shape or size.

A *liquid* has volume but no shape of its own. It is mobile and conforms to the shape of the containing vessel. It offers but slight resistance to any stress except one tending to change its volume.

A *gas* has the distinctive property of indefinite expansion. It has neither independent form nor independent volume, but completely fills the vessel containing it.

The term *fluid* applies both to liquids and gases.

While solids, liquids, and gases may be distinguished as above, there are nevertheless some substances which are neither wholly in the one state nor the other. Sealing wax passes imperceptibly by the aid of heat from the solid state to the liquid. Shoemaker's wax so far resembles a solid that it will break into fragments under the blow of a hammer, and yet under long-continued stress it flows like a liquid and can be molded at will.

In addition there is "the critical state," in which the properties exhibited by a substance do not determine conclusively whether it is a liquid or a gas.

16. The Constitution of Matter. — The theory that matter is not infinitely divisible is one of great antiquity. According to this theory, matter is made up of indivisible parts called

atoms. The term *molecule* has long been applied to the smallest subdivision of a substance which can exist by itself and still exhibit the properties of that substance. Thus a molecule of water is composed of a group of atoms, one of oxygen and two of hydrogen. When this group is broken up, the parts are distinct atoms of oxygen and hydrogen gas, and not water. This theory of the constitution of matter has prevailed for a very long period, and has served useful purposes in science.

If this theory were the ultimate truth, then the molecule would not be divisible without loss of physical and chemical properties, and the atom of one kind of matter could not be resolved into atoms of any other kind. The so-called "elements" of chemistry would then be the ultimate constituents of matter.

The phenomena of electrical discharges through rarefied gases have shown that a molecule may have dislodged from it very minute particles, or carriers of negative electricity, called *electrons;* and that these electrons have a mass only about $\frac{1}{1800}$ that of the hydrogen atom.

Moreover, the investigation of radium has led to the discovery that it disintegrates spontaneously into simpler and lighter atoms of other elements, notably helium. Hence it now appears entirely plausible that there is but one kind of "stuff" in the physical universe, perhaps hydrogen, and that all other kinds are fashioned from this ultimate element. While therefore atoms and molecules still exist, the concepts attaching to these terms are now very different from those which have prevailed for a long time.

III. PHYSICAL MEASUREMENTS

17. Units.—Every physical quantity considered in modern Physics has a definite magnitude, and to measure it a certain fixed amount of the same kind of physical quantity is employed as the *unit* of measurement. Thus, to measure

a length, some standard length, such as a foot, is taken as the unit, and the process of measurement consists in finding how many times this unit is contained in the given length.

The expression for the magnitude of any physical quantity always consists of two parts. One of these is the name of a certain quantity of the same kind as the quantity to be measured, which is taken as the standard or unit; the other is merely numerical and expresses the number of times the unit must be applied to make up the quantity measured. For example, (150) (feet), (50) (grams), (30) (seconds).

Further, the numerical parts of two expressions for the same quantity in different units are inversely as the magnitudes of the units employed. Thus, a certain smokestack is either 150 feet high, or 50 yards high; and since the yard is three times the foot, the numeric in yards is one third as great as in feet.

Since every physical quantity must be measured in terms of a unit of its own kind, there are as many units as there are different kinds of physical quantities to be measured.

18. Fundamental and Derived Units.—It has been found that nearly all the units for physical measurement may be defined in terms of three others, which are arbitrarily chosen. The three generally employed for the purpose are the units of *length*, *mass*, and *time*. These are called *fundamental units* to distinguish them from all others, which are called *derived units*. The system almost universally used in physical investigations employs the *centimeter* for the unit of length, the *gram* for the unit of mass, and the *second* for the unit of time. This system is accordingly known as the *c. g. s.* (centimeter-gram-second) system. The centimeter, the gram, and the second are then the fundamental units; the square centimeter as the unit of area, the cubic centimeter as the unit of volume, and the centimeter per second as the unit of velocity are examples of derived units.

19. Units of Length.—In the metric system the standard unit of length is the *meter* (m.). It is the distance between two transverse lines on a bar of platinum-iridium, at 0° C. (centigrade scale), constructed by the International Metric Commission and preserved in the vaults of the International Bureau of Weights and Measures at Sevres, a suburb of Paris. This is called the international prototype. The *centimeter* (cm.) is the one hundredth part of the meter. The only multiple of the meter in common use is the kilometer (km.), equal to 1000 meters. It is the unit employed on the continent of Europe for such distances as we express in miles. One kilometer is equal to 0.6214 mile.

The United States possesses two national prototype standard meters constructed by the International Commission and preserved at the Bureau of Standards in Washington.

By Act of Congress in 1866, the use of the metric system of weights and measures became lawful in the United States, and the weights and measures in common use were defined in terms of those of the metric system. By this same act the legal value of the yard in the United States is $\frac{3600}{3937}$ of a meter; conversely the meter is equal to 39.37 inches. The inch is 2.540 cm.

The unit of length in the English system for the United States is the *yard*, defined as above. In Great Britain it is the distance between the transverse lines in two gold plugs in a certain bronze bar at 62° F., preserved in the Standards' office, Westminster. One third of the yard is the *foot*, and one thirty-sixth is the *inch*.

20. Units of Mass.—The unit of mass in the metric system is the *kilogram*. It is the mass of the international prototype kilogram preserved with the international prototype meter at Sevres. The *gram* (gm.) is the one thousandth part of the kilogram. The kilogram (kgm.), was originally designed to represent the mass of a cubic decimeter (liter) of pure water at 4° C., the temperature of

the greatest density of water. For practical purposes this is the mass of a kilogram; and the gram is the mass of a cubic centimeter of water at the same temperature.

Two of the national prototypes, constructed of platinum-iridium by the International Metric Commission, are the standard kilograms for the United States. They are also preserved at the Bureau of Standards in Washington.

The standard unit of mass in the English system is the *avoirdupois pound.* The *ton* of 2000 pounds is the chief multiple in the United States; its submultiples are the *ounce* and the *grain.* The avoirdupois pound is equal to 16 ounces, and to 7000 grains. The coinage of the United States is regulated by the "troy pound of the mint," containing 5760 grains. By the law of May 16, 1866, the weight of the 5-cent nickel-copper piece was fixed at 5 grams; and by the law of February 12, 1873, the weight of the silver half dollar was fixed at 12.5 grams.

In accordance with the International Postal Convention, the metric system of weights was "adopted for international postal relations to the exclusion of every other system." The revised statutes of the United States for 1872 contain the clause that "fifteen grams (of the metric system) shall be the equivalent for postal purposes of one half ounce avoirdupois." The interchange of mail by all civilized countries represents the most extensive use of a uniform system of weights in the world.

21. The Unit of Time. — The unit of time in universal use is the *second* of mean solar time. An apparent solar, or sundial, day is the interval between two successive transits of the sun's center across the meridian of any place. But the apparent solar day varies in length from day to day by reason of the varying speed of the earth in its orbit and the inclination of its axis to the orbit. Hence the average length of all the apparent days throughout the year is taken as the length of a mean solar day. This day is divided into 86,400 equal parts, each of which is a second of mean solar time.

Mean solar or clock time agrees with sundial time on April 15, June 14, September 1, and December 24. At other times the difference be-

tween noon as indicated by a mean-time clock and a sundial is called the *equation of time*. It is the correction which must be applied to apparent time to get mean time. The maximum value of this equation of time may amount to plus 14 min. 32 sec. and minus 16 min. 18 sec., the dates and

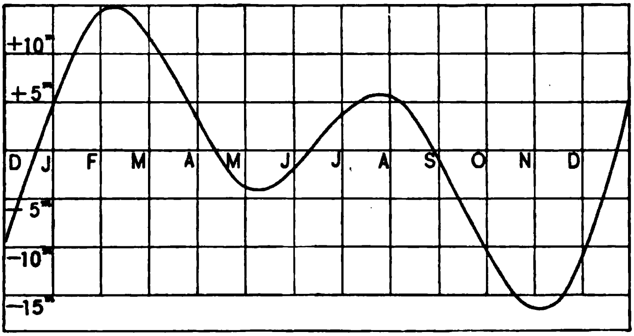


Fig. 3

amounts varying slightly from year to year. The curve in Figure 3 represents the equation of time for the year. The positive ordinates of the curve mean that the mean-time clock is ahead of the transit of the sun.

The astronomical unit of time is the sidereal day. It is the interval between two successive transits of a star across the meridian. Since the diameter of the earth's orbit is small compared with the distance of a fixed star, the line joining the earth and a star always remains parallel to itself. Hence the sidereal day represents the period of the rotation of the earth on its axis. A sidereal day is equal to 23 hours, 56 minutes, 4.09 seconds of mean solar time; that is, the sidereal day is nearly 4 minutes shorter than the mean solar day.

22. Units of Angular Measure. — The unit of angular measure commonly used in numerical calculations is the *degree* of arc. The circumference of a circle corresponds to 360 degrees, and a right angle to 90 degrees. A degree is divided into 60 minutes of arc, and a minute into 60 seconds of arc. The degree of arc is a relic of an ancient sexagesimal system, consisting of multiples of 60.

The other system of measuring angles is the *circular measure*, in which the unit is the *radian*. The radian is the angle subtended by an arc equal in length to the radius of the circle. If the length of the arc is s and the radius of the circle r, the

angle subtended by this arc is s/r. If the radius is unity, then the length of the arc measures the subtended angle in radians. Hence in circular measure $\pi/2$ is equivalent to 90°, π to 180°, and 2π to 360°. Since

$$2\pi \text{ radians} = 360°,\ 1 \text{ radian} = 360°/2\pi = 57.2958°.$$

23. Trigonometrical Functions. — The expression of physical relations quantitatively is greatly facilitated by the use of a few simple trigonometrical relations. It will be convenient to define them here for students unfamiliar with the elements of plane trigonometry.

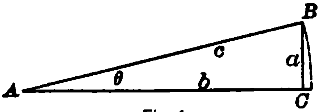


Fig. 4

Let ABC (Fig. 4) be a right triangle, and let its sides be denoted by the letters a, b, c. Then the following terms may be defined as ratios:

$$\text{sine } \theta = \frac{a}{c};\ \text{cosine } \theta = \frac{b}{c};\ \text{tangent } \theta = \frac{a}{b}.$$

It follows that $a = c \sin\theta$; $b = c\cos\theta$; $a = b\tan\theta$.

Also $a^2 + b^2 = c^2(\sin^2\theta + \cos^2\theta)$. And from the right triangle $a^2 + b^2 = c^2$.

Therefore, $$\sin^2\theta + \cos^2\theta = 1.$$

Also from the first definitions,

$$\frac{\sin\theta}{\cos\theta} = \frac{a}{b} = \tan\theta.$$

When the angle θ is very small, the side a may be identified with the arc subtending the angle, and the side b with the radius c.

Then $$\sin\theta = \frac{a}{c} = \theta;\ \tan\theta = \frac{a}{b} = \theta;\ \cos\theta = \frac{b}{c} = 1.$$

For small angles, the angle in radians is equal to either the sine or the tangent of the angle, and the cosine is unity. These relations will be found of considerable utility in simplifying expressions involving these trigonometrical functions.

$4 = \frac{8}{2}$ $\quad 8 \div 2$

CHAPTER II

KINEMATICS

I. RECTILINEAR MOTION

24. Motion of a Material Particle.—A *material particle* is an ideal body assumed to be without sensible dimensions. The limiting size of a material particle is relative only. In some astronomical problems, the earth and the planets may be treated as material particles, since their dimensions are negligible on a scale representing their distances from the sun; on the other hand, in studying the phenomena of electric discharges through rarefied gases, it may not be permissible to regard the molecule even as a material particle.

If a material particle has relatively no sensible dimensions, its position may be denoted by a geometrical point. Now when such a particle is displaced from point to point, it must occupy in succession every point along the path of displacement, and time must be consumed in the operation. *Motion* is the process of change in the position of a material particle, considered as taking place during a definite interval of time.

Kinematics is the science of the motion of a material particle without reference to its cause or to any physical actions that produce changes of motion.

25. Types of Motion.—In *motion of translation* any line joining two points in a body maintains the same direction; that is to say, its successive positions are all parallel to one another. In addition to motion of translation, a body may spin or rotate about a definite line in the body, and this

motion is called *rotatory motion.* A material particle is capable of motion of translation only; or, at least, its rotation is of no dynamical significance. In general the motion of an extended body is a combination of a rotation and a translation.

Neglecting the curvature of the earth, a steamship sailing in a straight line is an example of motion of translation. The motion of the armature and pulley of a stationary electric motor is one of pure rotation. The wheel of a moving locomotive and a ball rolling along the floor of a bowling alley combine motion of rotation with motion of translation.

The translatory motion of a material particle may be either *rectilinear* or *curvilinear.* The present section is restricted to the subject of rectilinear motion, or motion along a straight line.

26. Velocity Uniform and Variable. — *Velocity is the time rate of motion of a body.* By the time rate of motion is meant the distance traversed in a given time divided by that time. If a point moves over equal spaces in equal successive time intervals, its motion is *uniform* and its velocity *constant.* The motion of a star across the field of view of a fixed telescope is an instance of uniform motion. The speed of a railway train may be constant for a considerable distance. It may, for example, travel 88 feet for each second of an entire minute, or at the rate of a mile a minute.

When a body traverses unequal spaces in successive equal periods, its motion is *variable.* The motion of a falling body is variable, for it moves faster and faster as it descends. The velocity at any instant in variable motion is the distance the body would move in the next unit of time if at that instant its motion were to become uniform without other change. For example, the velocity of a shell as it leaves the muzzle of a gun is the distance it would pass over in the next second if it should continue to move uniformly without disturbance. The velocity of a falling body at any instant is the distance it would fall during the next second

from that instant, if the attraction of the earth and the resistance of the air could be withdrawn.

27. Formulæ for Uniform Motion.—Let v be the constant velocity of a body moving with uniform motion. Then, if the space s is passed over in t units of time, the velocity will be given by the relation

$$v = \frac{s}{t}. \tag{1}$$

From the same relation we have $s = vt$ and $t = \frac{s}{v}$.

Even though the motion is not uniform, if the space s is passed over in time t, the *mean* or *average velocity* is still given by equation (1). If both the space and the time be reduced to indefinitely small quantities, then the mean velocity becomes the actual velocity for the instant.

The practical unit of velocity in the *c. g. s.* system is the velocity of *one centimeter per second.*

28. Acceleration.—*Acceleration is the time rate of change of velocity.* A case of special interest is one in which the change in velocity is the same from second to second. The motion is then one of *uniform acceleration.* If the velocity increases, the acceleration is positive; if it decreases, it is negative. When a heavy body falls, its gain in *velocity per second* is 9.8 m. *for every second it falls.* Its acceleration is, therefore, 9.8 m. per second per second; in other words, an increase in velocity of 9.8 m. per second is acquired in a second of time. This acceleration is the same as an increase in velocity of 588 m. per second acquired in a minute of time.

If a railway train should start from rest and increase its speed one foot a second for a whole minute, its velocity at the end of the minute would be 60 feet a second. Since it would acquire in one second a velocity of one foot a second, and in one minute a velocity of 60 feet a second, its acceleration would be either one foot per second per second, or 60 feet per second per minute. Acceleration is expressed in terms of the fundamental units of length and time, the first power of a length and the negative second power of a time.

Let v_0 be the initial velocity at the instant from which the time is counted, v the final velocity at the time t, and a the acceleration. Then by definition

$$a = \frac{v - v_0}{t}. \tag{2}$$

From this equation $v = v_0 + at.$ (3)

The practical unit of acceleration is the acceleration of one centimeter per second per second in the *c. g. s.* system, or one foot per second per second in the English system.

29. Formulæ for Uniformly Accelerated Motion. — Since the gain in velocity is constant, the average velocity is one half the sum of the initial and final velocities, and the space s passed over is the product of this average velocity and the time, or

$$s = \frac{v + v_0}{2} t. \tag{4}$$

Substitute for v its value from equation (3), and

$$s = \frac{2 v_0 + at}{2} t = v_0 t + \tfrac{1}{2} at^2. \tag{5}$$

Multiply together (2) and (4), and

$$as = \frac{v^2 - v_0^2}{2},$$

or

$$v^2 = v_0^2 + 2 as. \tag{6}$$

If the initial velocity is zero, (3), (5), and (6) become

$$v = at, \tag{7}$$

$$s = \tfrac{1}{2} at^2, \tag{8}$$

$$v^2 = 2 as. \tag{9}$$

30. Scalar and Vector Quantities. — A *scalar quantity* is one having magnitude only. A complete specification of scalar quantities is made by giving their numerical value in terms of the proper unit. Thus, 100 cm.³ of water, 10 kgm. of

sugar, 50 minutes of time, are all completely expressed scalar quantities. Volume, mass, time, density, energy, etc., are scalar quantities.

A *vector quantity* has not only magnitude but direction, and a vector quantity is not completely expressed unless its direction is given as well as its magnitude. The difference between scalar and vector quantities becomes apparent whenever it is necessary to add together two or more vectors. Thus, if a steamship is propelled by its screw 20 miles an hour, and at the same time is driven by the wind 5 miles an hour, the distance actually traveled in an hour is indeterminate unless both directions are given in which the ship is driven. If one railway porter pulls on a truck with a force of 200 units and another with a force of 150 units, the total force applied to move the truck is indeterminate, both in magnitude and direction, unless the directions of both pulls are given. Displacement, velocity, acceleration, momentum, force, etc., are vector quantities.

The addition of scalar quantities is effected by simple arithmetic; the addition of vector quantities involves their directions and is effected by geometrical operations. It is of great advantage to represent vector quantities by straight lines, the length of the line denoting the magnitude of the vector, and the direction in which it is drawn, the direction of the vector.

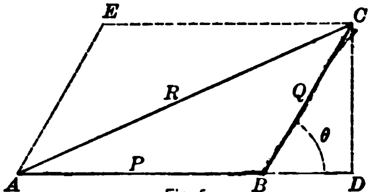


Fig. 5

31. Addition of Vectors. — Let two vector quantities, P and Q, be represented by the lines AB and BC (Fig. 5). A particle may suffer, for example, two successive displacements, the first represented by the line AB, and the second by the line BC. Then the resultant displacement R, or the single displacement which would leave the point in the same position as the two successive displacements, is represented by the line AC.

Or further, suppose the point to have two velocities at the same time, one represented by AB, and the other by AE. The actual resultant velocity is represented by the diagonal AC; for if the point moves uniformly along the line AB in one second, and at the same time AB is carried parallel to itself the distance BC, the point B moving to C, then the actual path of the moving point is AC. Hence the vector sum of P and Q is represented by the third side R of the triangle ABC. The triangle is half of the corresponding parallelogram. Hence the following parallelogram law:

If two vectors, which are applied to a material particle at the same time, are represented in magnitude and direction by two adjacent sides of a parallelogram, then their vector sum will be represented in magnitude and direction by the diagonal of the parallelogram drawn through the intersection of these two sides.

The magnitude of the resultant vector may be calculated from Figure 5 as follows: Let θ be the angle between the vectors, that is, the angle through which Q must be turned to coincide with P. Then

$$AC^2 = AD^2 + DC^2,$$

or

$$R^2 = (P + Q\cos\theta)^2 + Q^2\sin^2\theta.$$

$$= P^2 + 2PQ\cos\theta + Q^2 \quad (\S\ 23). \tag{10}$$

This formula applies to all values of θ. Thus, if $\theta = 0$, $\cos\theta = 1$, and $R = P + Q$; if $\theta = 90°$, $\cos\theta = 0$, and $R^2 = P^2 + Q^2$; if $\theta = 180°$, $\cos\theta = -1$, and $R = P - Q$.

32. Resolution of Vectors. — A vector may be resolved into components in any given directions. This operation is the reverse of finding the vector sum. The given vector is assumed to be replaced by two or more component vectors, which are so chosen that the actual vector is the resultant of these assumed vectors.

The component vector in any given direction is its value as a vector in that direction. The most common case is

the resolution of a vector into two components in two directions at right angles to each other. In most cases it suffices to find the component vector in the direction in which the attention for the time being is directed; the other component at right angles to the first is obviously without effect in the direction considered.

To illustrate: Let it be required to resolve a velocity of 8 m. a second into two rectangular components.

First. Suppose one of the components is to be 3 m. per second; find the other component. The problem is to construct a rectangle with a diagonal of 8 and one side 3, to find the adjacent side.

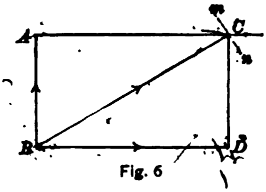


Fig. 6

Draw BA (Fig. 6) 3 units in length, and at A draw AC perpendicular to BA. With B as a center and with a radius of 8 units, draw the arc mn cutting AC at C. Complete the rectangle $ABDC$. Then BA and BD are the two components of BC, and BD is the one required. Its value may be found from the right triangle BDC.

$$BD = \sqrt{8^2 - 3^2} = 7.416.$$

Second. It is required to find the component BD in a direction making an angle θ with the vector BC. From § 23,

$$BD = BC \cos\theta.$$

If BC and the direction angle θ are given, the problem is easily solved numerically. In general, *a component vector in any direction (BD) is found by multiplying the given vector (BC) by the cosine of the direction angle (θ).*

II. CURVILINEAR MOTION

33. Uniform Circular Motion.—Up to this point a velocity has been assumed to vary in magnitude only, and therefore the acceleration has been confined to the direction of motion.

But a velocity may vary in direction as well as in magnitude. If a particle has a uniform motion in a straight line, its acceleration is zero in every direction. If its velocity changes in magnitude only, then its acceleration is positive or negative along the line of motion. But if the *direction of motion* changes, then the particle has at least a component acceleration at right angles to its path, and its motion is *curvilinear.*

For example, if a particle moves uniformly along AB (Fig. 7), while from B to C it describes a curved path, then between these two points there is an acceleration normal to the path.

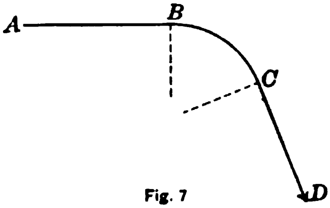


Fig. 7

In uniform circular motion, the velocity of the particle along the circumference of the circle is constant; but since its rate of deflection from the tangent to the circle is constant, the acceleration is constant and is directed everywhere toward the center of the circle. If it were not toward the center, it could be resolved into two components, one toward the center and the other along a tangent to the circle; the latter would mean a change of velocity in the circle itself. But the velocity in the circle is uniform, and there is therefore no tangential component of acceleration, or the acceleration is directed wholly toward the center. In other words it is centripetal.

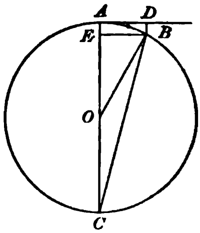


Fig. 8

34. Centripetal Acceleration. — Let ABC (Fig. 8) be the circle in which the point revolves, and AB the very small portion of the circular path described in the time t. Denote the length of the arc AB by s. Then, since the motion is uniform, $s = vt$.

AB is the diagonal of a very small parallelogram $AEBD$, and AE is the distance through which the moving point

is deflected toward the center while traversing the small distance AB. Since the acceleration is constant, we have, from equation (8) for uniformly accelerated motion,

$$AE = \tfrac{1}{2} at^2.$$

The two right triangles ABE and ABC are similar, and therefore $$AE : AB = AB : AC.$$
Whence $$AB^2 = AE \times AC.$$

Substitute in this equation the values of AB and AE above, and for AC the diameter $2r$, and the equation becomes

$$v^2t^2 = \tfrac{1}{2} at^2 \times 2r = at^2r;$$

whence $$a = \frac{v^2}{r}. \tag{11}$$

If T be the period of rotation of the point,

$$v = \frac{2\pi r}{T} \quad \text{and} \quad v^2 = \frac{4\pi^2r^2}{T^2}.$$

Substitute this value of v^2 in (11), and finally

$$a = \frac{4\pi^2r}{T^2}. \tag{12}$$

III. PERIODIC MOTION

35. Definition of Periodic Motion. — When a body goes through the same series of movements at regularly recurring intervals, its motion is said to be *periodic*. Thus the motion of the earth in its orbit around the sun is periodic. Its velocity is not uniform, but at intervals of a year it returns to the same value in the same direction. If the motion returns periodically to the same value, and in addition is periodically reversed in direction, it is then *vibratory* or *oscillatory*. The motion of a pendulum, of a violin string, or of the prong of a tuning fork, is vibratory.

36. Simple Harmonic Motion.—*Simple harmonic motion* is the projection of uniform motion in a circle, either on a diameter or on a line in the plane of the circle.

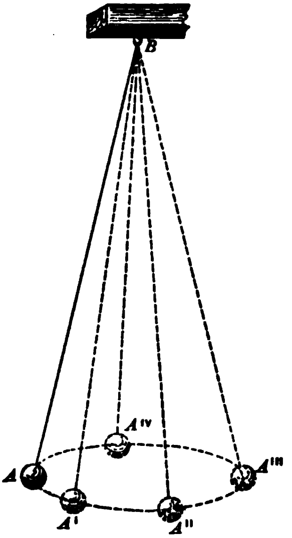


Fig. 9

Suspend a ball by a long thread and set it swinging in a small horizontal circle (Fig. 9). It will travel round and round with uniform speed, the thread describing the surface of a cone. Place a white screen back of the ball; stand a few feet away, and with the eye on a level with the ball, watch the projection of the ball on the screen. The eye discerns the motion to the right and left of the line of sight, but not the motion toward the observer and away from him. The apparent motion of the ball is simple harmonic, or like the motion of a simple pendulum.

Let the circle of Figure 10 represent the path of the ball, and *ABCD*, etc., its projection on the screen. When the ball moves over the arc *adg*, it appears to an observer far to the left of the figure to be moving from *A* through *B*, *C*, etc., to *G*, where it momentarily comes to rest. It then starts back toward *A*, at first very slowly, but with increasing velocity till it passes *D*. Its velocity then diminishes, and at *A* is again zero and a reversal of the motion takes place. At *a* and *g* the actual motion of the ball is all in the line of sight,—away from the observer at *a* and toward him at *g*. At *k* and *d* the ball is moving across the line of sight and the projected motion at *D* is the fastest.

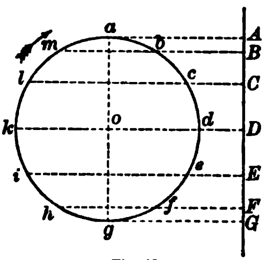


Fig. 10

When a point vibrates to and fro along a straight line, as *AG*, in such a manner that its position at any moment is the same as the projection on that line of a point moving uniformly in a circle whose diameter is the length of the straight line, its motion is simple harmonic.

The circle *adgk* is called the *circle of reference*. Its radius is the *amplitude* of vibration.

The *period* of the motion is the time of a double oscillation, or the time of a complete revolution of the auxiliary point around the circle of reference. The reciprocal of the period, that is, the number of vibrations per second, is called the *frequency*. For example, if the period is $\frac{1}{10}$ of a second, the frequency is 10 double vibrations per second.

Motion from left to right is *positive*, and from right to left *negative*. Displacement to the right of the middle point is *positive*, and to the left *negative*.

The *phase* is the fraction of a period which has elapsed since the particle last passed through the middle point of its path in the positive direction.

37. The Harmonic Curve. — *The harmonic curve*, or "curve of sines," is the result of combining a simple harmonic motion (s. h. m.) with a uniform motion at right angles to it. In

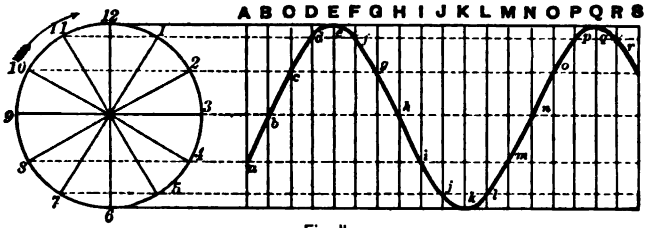


Fig. 11

Figure 11 the equal distances between the vertical lines *A*, *B*, *C*, etc., represent the uniform motion along a horizontal line. The circle is the circle of reference with a radius equal to the amplitude of vibration. Its circumference is divided into any convenient number of equal parts (some multiple of four

is best), and through the points of division are drawn the horizontal lines cutting the vertical ones. The horizontal lines divide the verticals into spaces traversed by the particle in successive twelfths of a period.

If the particle is in the line A for the horizontal motion, and in the horizontal line through 8 on the circle of reference for the s. h. m. motion, it must be at the intersection of the two at a. After one twelfth of a period it will be at b, etc. The desired curve is found by drawing a smooth curve through the successive intersections of the two sets of lines.

Such a curve may be drawn experimentally by causing a large tuning fork to inscribe its vibrations on smoked paper fastened around a drum, which is rotated with uniform angular velocity, while a light tracing point attached to the fork inscribes a sine curve.

It may also be drawn in a very simple way by means of a long flat strip of clear wood, securely mounted horizontally by one end so as to vibrate in a horizontal plane, and carrying at the other end a small camel's-hair brush saturated with ink. A long strip of paper attached to a narrow board is drawn as uniformly as possible against a guide parallel to the vibrating wood strip, and the brush marks the harmonic curve. The zero line is drawn in the same manner with the brush at rest.

38. Acceleration in Simple Harmonic Motion.—Let the auxiliary point be at B (Fig. 12) in the circle of reference and moving toward Y. The corresponding displacement of the point executing simple harmonic motion along the diameter through X is OC; denote it by x. The only acceleration present is the centripetal acceleration a along the radius BO. The problem is to find the component of this acceleration in the direction of the harmonic motion along OX. Denote this component by a_x. Acceleration is a vector, and by § 32, the component of a

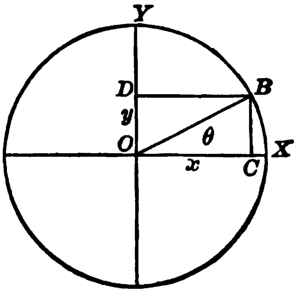


Fig. 12

vector in any direction is found by multiplying by the cosine of the direction angle θ. Therefore

$$a_x = -a \cos \theta.$$

But $\cos \theta = \frac{x}{r}\cdot$ Hence $a_x = -\frac{a}{r}\cdot x = -\frac{4\pi^2}{T^2}\cdot x.$

Since both a and r are constants, *the acceleration along OX is proportional to the displacement* x, and opposite in sense.

The acceleration in simple harmonic motion is always directed toward the middle point of the path of the moving point; and the proportionality of the acceleration to the displacement is the distinguishing characteristic of simple harmonic motion.

Similarly the acceleration parallel to OY is $-\frac{a}{r}\cdot y$. The two accelerations differ in phase by a quarter of a period, for when a_x is a maximum, a_y is zero, and conversely. Obviously *uniform circular motion is composed of two simple harmonic motions at right angles to each other, of the same period and amplitude, and differing in phase by a quarter of a period.* If the amplitudes are not equal, the resulting motion is in an ellipse.

To illustrate the composition at right angles of two oscillations of the same period, suspend a steel ball by a long, fine fishline so that the ball just clears the surface of a table. Set it swinging north and south, and strike it with a block of wood in an east and west line as it crosses the middle point of its swing. The resultant oscillation will be in a diagonal line between the north-south and east-west directions. The two oscillations combined are in the same phase.

Start the ball as before and strike it at right angles to its path as it reaches the extreme limit of its swing. If the blow is rightly gauged, the resulting motion will be sensibly circular. The two motions combined are nearly simple harmonic, of the same period and amplitude, and they differ in phase by a quarter of a period.

39. Velocity in Simple Harmonic Motion. — The acceleration in simple harmonic motion is greatest when the displacement is greatest, or at either limit of the swing. It declines from that point to the middle position, where it

becomes zero. The velocity, on the other hand, is zero at either limit of the excursion of the oscillating particle and increases from that point to a maximum at the median position. The point starts from rest with the greatest acceleration, or rate of change of velocity, and its velocity increases all the way till it passes the middle point, *but at a constantly decreasing rate.* The gain in velocity for each equal increment of time is less and less as the point approaches the middle of its excursion, but becomes zero only as the point passes this median position, after which the velocity decreases at a constantly increasing rate up to rest at the other limit of displacement.

Acceleration and velocity in simple harmonic motion stand in such a relation that when one is greatest the other is least. At either limit of motion, the acceleration is the same as the centripetal acceleration in the circle of reference, while the velocity is zero; *at the middle point of the motion, the velocity is the same as in the circle of reference*, while the acceleration is zero. If, as in the last article, the acceleration is proportional to $\cos\theta$, the velocity is proportional to $\sin\theta$, since the sine and cosine of an angle are related as described above for acceleration and velocity in simple harmonic motion.

40. Composition of Two Simple Harmonic Motions in the Same Direction. — *A. When the periods are the same:* The simplest case of the composition of two s. h. m.'s along the same line occurs when the two motions have the same period. They may differ in phase and amplitude, but their resultant will always be simple harmonic and of the same period as that of the component motions. Their composition is readily effected graphically by means of the harmonic curve of § 37.

Let $ABCDE$ (Fig. 13) be the harmonic curve corresponding to one s. h. m., the amplitude being bB and the period AE; and let $abcde$ represent the other s. h. m. of amplitude Cc and of the same period as the first. Further, the second

s. h. m. is one quarter of a period *behind* the first in phase. Then the resulting displacements will be found by adding together the corresponding displacements of the two s. h. m.'s with their proper signs. Thus the resultant displacement at

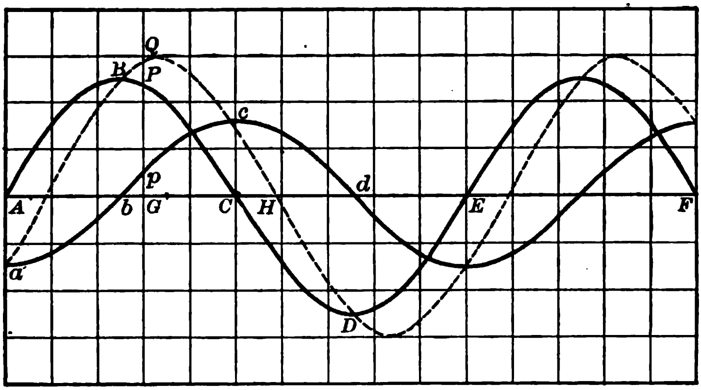


Fig. 13

the instant denoted by the point G is the sum of Gp and GP or GQ. At C it is Cc, since the displacement of the curve $ABCDE$ is zero. At H the two displacements are equal and of opposite sign; the resultant is therefore zero. The

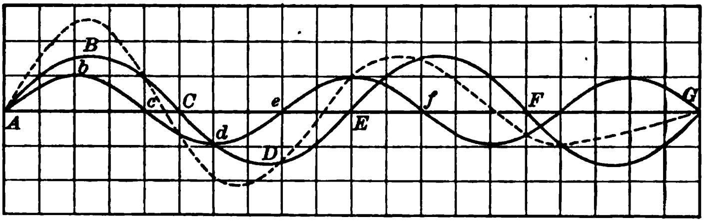


Fig. 14

dotted curve is the resultant due to the composition of the other two. It is harmonic, of the same period as the component curves, and intermediate in phase between them.

B. When the periods are not the same. If the periods of the two s. h. m.'s combined are not the same, the resultant

curve will be periodic, but not harmonic. Moreover, its period will be longer than that of either of the combined s. h. m.'s. Let *ABCDE* (Fig. 14) be the harmonic curve corresponding to one of the s. h. m.'s and *Abcde* to another of different period and amplitude. The period of the first is *AE* and that of the second is *Ae*. The resulting displacements are obtained in the same manner as before by adding corresponding ordinates with their proper signs. The dotted curve is the result. Its ordinates are everywhere equal to the algebraic sum of the corresponding ordinates of the two component harmonic curves. The figure shows only half of the complete periodic curve. The other half is similar to this, taken in the reverse order, and all ordinates of opposite sign.

The same harmonic curves are combined in Figure 15, but the scale is smaller, so that four periods of the one and five

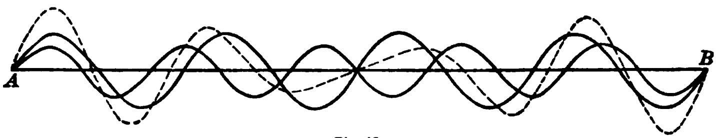


Fig. 15

of the other are included in the diagram. This gives the complete resultant periodic curve. Its period is four times that of the slower oscillation and five times that of the quicker one.

The composition of two s. h. m.'s of very nearly the same period in the same direction gives rise to a periodic curve which is everywhere very nearly a sine curve in form, but its amplitude increases and decreases periodically. The maxima occur when the component vibrations are the same in phase, and the minima when they are opposite in phase. From maximum to maximum one vibration gains exactly one period on the other. If, for example, the frequencies are as 24 to 25, a maximum occurs at every 25th period of the quicker vibration.

The waxing and waning of the resultant amplitude when two s. h. m.'s of nearly equal periods are combined in the same direction explain the familiar phenomenon of beats in music. This will be illustrated in the part on Sound.

41. Composition of Two Simple Harmonic Motions at Right Angles.—The rectangular composition of two s. h. m.'s is easily effected graphically by means of two circles of reference. Their periods may be the same, or they may be in any simple ratio.

A. Periods equal. Let the radii of the two half circles *ADG* and *adg* (Fig. 16) be the relative amplitudes of the two s. h. m.'s. Divide the two circumferences into the same number of equal parts, and through the points of division draw parallel horizontal and vertical lines respectively. The rectangle, whose sides are *LM* and *LN*, will contain all the figures resulting from the composition of the two s. h. m.'s with any phase difference.

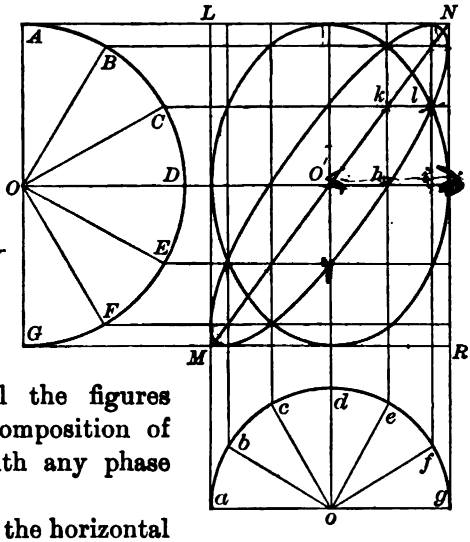


Fig. 16

$\frac{1}{12}$ T

The spaces between the horizontal lines represent the distance traveled in equal fractions of a period for the s. h. m. of larger amplitude, OA; and those between the vertical lines represent the distances traveled in the same equal intervals by the s. h. m. of smaller amplitude, oa. The intersection O' of the lines through the centers of the circles corresponds to a phase difference of zero between the two component motions. Each point of intersection along $O'j$ corresponds in this figure to a phase difference of one twelfth of a period. Thus, if i is one point of the resultant curve, the difference of phase is one sixth of a period; if j is on the resultant curve, the phase difference is one quarter of a period.

If we start with no phase difference, the moving point will be at O' for both the horizontal and vertical motions. After one twelfth of a period, it will have advanced one division to the right and one upward, and the point satisfying both these conditions is the opposite corner k of the small rectangle. Continuing in this way, it will be found that the point traces the diagonal MN of the large rectangle.

If the phase difference is one quarter of a period, the starting point is j. One space to the left and one up gives l for the position after one twelfth of a period. Continuing in this way and passing a smooth curve through the successive diagonal corners of the small rectangles, the resultant is the large ellipse, whose major and minor axes are the diameters of the two circles of reference. If at the same time the two amplitudes are equal to each other, the ellipse becomes a circle (§ 38).

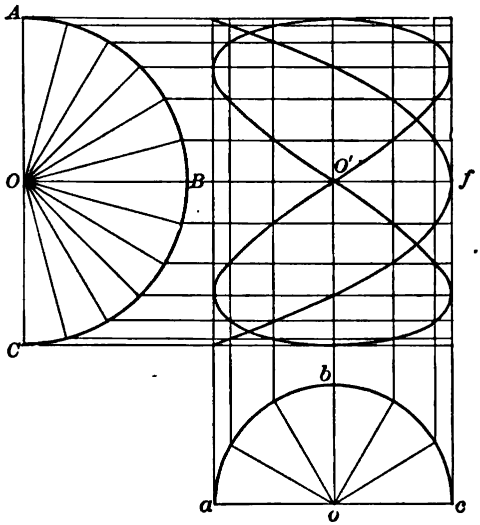


Fig. 17

If the phase difference is one twelfth of a period, the resultant motion is in the smaller ellipse. If the periods of the two s. h. m.'s are not exactly equal, the resultant curve will pass through all the possible ellipses in succession between the two diagonal straight lines of the large rectangle as limits, the large ellipse being the intermediate form. During the passage from one of these diagonals over to the

other and back again, one of the component motions has gained a complete vibration on the other.

B. Periods one to two. In Figure 17 the larger circle has twice as many divisions as the smaller one. If the spaces between the vertical and horizontal lines denote the distances traveled in equal successive intervals of time for the two motions respectively, then the periods of the s. h. m.'s combined are as one to two.

With no phase difference, the resulting curve is the figure 8 (the lemniscate), traced in the same manner as already described. If the phase difference is a quarter of a period, the resultant curve is the parabola through *f*. For a phase difference of three quarters of a period, the curve will be a parabola with its vertex on the opposite side of the rectangle. If the periods are not rigorously as one to two, the two parabolas are the limiting forms of the resultant curve, and the figure 8 is intermediate between them.

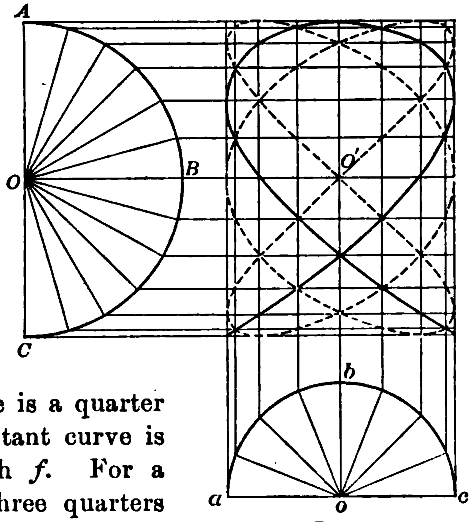


Fig. 18

C. Periods two to three. For this case the number of equal divisions of the circles of reference are as two to three (Fig. 18). By changing the phase difference one eighth of a period, the resultant curve passes from the full line figure to the dotted one.

If in tracing the curve we count the passages across the large rectangle, both vertically and horizontally, we can always determine the ratio of the periods combined. Thus, the full line curve of Figure 18 is completed by two passages

across vertically, and by three passages across horizontally. The two periods combined are therefore as two to three.

Problems

1. A railway train has a speed of 60 mi. an hour. What is its speed in feet per second?

2. A body moving uniformly in a circular path of 15 m. radius makes 10 revolutions in 5 sec. Find the speed per second.

3. If the radius of the earth is 4000 mi., what is the linear velocity per second of a point on the equator due to the earth's rotation on its axis?

4. A railway train 430 m. long passes over a bridge 150 m. long at a speed of 45 km. an hour. How long a time does the train take to pass completely over the bridge?

5. A body starting from rest passes over a distance of 256 m. in 4 sec. What is its acceleration?

6. A body starts with an initial velocity of 300 m. per second. If it comes to rest in 1 min. 2.5 sec., find the uniform negative acceleration.

7. A body has an initial velocity of 6 m. per second. Find its velocity at the end of 3 and 6 sec. respectively, if a equals 9.8 m. per second per second.

8. A train running at 60 mi. an hour is stopped with uniform retardation in 44 sec. by the application of brakes. What is the retardation per second per second?

9. How far will a ball roll before coming to rest if its motion is uniformly retarded, its initial velocity being 15 m. per second and the duration of its motion 10 sec.?

10. A body is projected upward with any initial velocity, and t and t' denote the times during which it is respectively above and below the middle point of its path. Find the ratio t/t'.

11. What is the final velocity of a body which passes over a distance of 144 m. in 1 min. with uniform acceleration: (a) when the initial velocity is zero; (b) when the initial velocity is 10 cm. per second?

12. What must be the initial vertical velocity of a ball to return to its starting point in 8 sec.? (The acceleration of gravity is 980 cm. per second per second.)

13. What acceleration per minute per minute must a body have to acquire in 20 min. a speed of 20 mi. per hour?

14. At what angle with the shore must a boat be steered in order to reach a point on the other shore directly opposite, if the actual velocity of the boat directly across is 8 mi. an hour and that of the stream 4 mi. an hour?

15. A locomotive driving wheel is 2 m. in diameter; if it makes 200 revolutions per minute, what is the average linear velocity of a point on the periphery? What is its greatest velocity? What is its least?

16. How far will a body move along a horizontal plane from rest in 30 sec., if it has an acceleration of 3000 cm. per second per second?

17. What is the acceleration of gravity g where a body falls 485 cm. in the first second?

18. A body moves along a horizontal plane with an acceleration of 360 m. per minute per second. How far will it travel in the fourth second?

19. A body slides down a smooth inclined plane and has a velocity of 10 m. per second at the end of four seconds. How far will it slide in the next four seconds?

20. A body moves uniformly around a circle 40 cm. in diameter at the rate of 24 revolutions per minute. Compute the acceleration toward the center.

21. A body moves around a circle with uniform velocity once a second and its centripetal acceleration is 1974 cm. per second per second. What is the diameter of the circle?

22. A vector drawn east has a length of 30 cm. and one drawn northeast a length of 50 cm. Find their vector sum.

23. An elastic rod is clamped at one end; when the other end is pulled aside 1 cm. and then released, it starts with an acceleration of 4 cm. per second per second. What is the period of its vibration?

24. In the last problem what is the greatest velocity of the free end of the vibrating rod?

25. A man weighing 75 kgm. stands on the platform of an automatic weighing machine which is placed on the floor of an elevator. What will be the indicated weight of the man when the elevator starts to descend with an acceleration of 100 cm. per second per second, if the acceleration of gravity is 980 cm. per second per second?

CHAPTER III

DYNAMICS

I. NEWTON'S LAWS OF MOTION

42. Dynamics Defined. — Up to this point the motion of a body has been considered in the abstract; and although the motion has been assumed to vary in certain definite ways, no inquiry has been made into the cause of these variations. A further step must now be taken by making this inquiry. That branch of Mechanics which studies the effects of force in producing motion or change of motion of definite masses of matter is called *Dynamics*.

43. Momentum. — Before proceeding to a discussion of Newton's laws of motion, which outline the relations between force and motion, it is necessary to define two terms associated with these laws. One of them is *momentum*. It is the product of the mass and the linear velocity of a moving body.

$$\textit{Momentum} = \textit{mass} \times \textit{linear velocity}, \text{ or } M = mv. \qquad (13)$$

In the *c. g. s.* system, the unit of momentum is the momentum of a mass of 1 gm. moving with the velocity of 1 cm. per second.

The change in the momentum of a body is proportional to the change in its velocity, since its mass is a fixed quantity. Hence the rate of change of momentum is proportional to the rate of change of velocity, that is, to the acceleration.

44. Impulse. — In estimating the effect of a force, the time during which it acts and its magnitude are of equal importance. This effect is doubled if the magnitude of the force is doubled, or if the time it continues to act is doubled.

Impulse is the product of the magnitude of a force and the time it continues to act.

Suppose a ball of 15 gm. mass fired from a rifle with a velocity of 40,000 cm. a second. Its momentum would be 600,000 units. If a truck weighing 300 kgm. moves at the rate of 2 cm. a second, its momentum is also 600,000 units. The ball acquires its momentum in a fraction of a second, while force may have been applied to the truck for perhaps thirty seconds to give to it the same momentum. In some sense the effect of the force in giving motion to the ball is the same as that of the force required to give the equivalent motion to the truck, because their momenta are equal. This equivalence is expressed by the term *impulse*, which is the same in the two cases.

Forces of very short duration, like the blow of a hammer, were formerly called impulsive forces, and their effect an impulse. The term *impulse* is now used in the more general sense as the product of a force and the duration of its action.

45. Laws of Motion. — Newton's laws of motion are to be regarded as physical axioms, incapable of rigorous experimental proof. They are axiomatic to those who have sufficient knowledge of physical phenomena to interpret their relations. The laws of motion must be considered as resting on convictions drawn from observation and experiment in the domain of physics and astronomy.

The most powerful argument for the validity of the laws of motion rests on the fact that their application to the solution of problems in mechanics leads to results which always agree with those of observation. The time of a coming eclipse, for example, is calculated by assuming the truth of Newton's laws; and the remarkable agreement between the calculated time and that subsequently observed confirms the laws.

These laws as enunciated by Newton are:

I. *Every body continues in its state of rest or of uniform motion in a straight line, except in so far as it may be compelled by impressed force to change that state.*

II. *Change of motion is proportional to the impressed force, and takes place in the direction in which the force acts.*

III. *To every action there is always an equal and contrary reaction; or the mutual actions of two bodies are always equal and in opposite directions.*

By "*change of motion*" we are to understand *change of momentum*, and by "*impressed force*," *impulse.*

46. Discussion of the First Law. — This law is known as the law of *inertia*, since it states that a body persists in its condition, either of rest or uniform motion, unless it is compelled to change that state by the intervention of an external force. It is further true that a body offers resistance to any such change in proportion to its *mass*. Hence the term *mass* is now commonly used to denote the measure of the body's inertia.

From this law we derive a definition of force, for the law asserts that *force is the sole cause of change of motion.*

47. Discussion of the Second Law. — The first law teaches that a change of momentum is due to impressed force. The second law points out what the measure of this force is. Maxwell restated it so as to make it read as follows: "*The change of momentum of a body is numerically equal to the impulse which produces it, and is in the same direction.*" By a proper choice of units, impulse may be placed equal to the change of momentum produced, or

$$Ft = mv. \tag{14}$$

Hence

$$F = \frac{mv}{t} = ma. \tag{15}$$

The initial velocity at the instant when the force begins to act is here assumed to be zero, and the final velocity at the conclusion of the force action is v ; mv is therefore the change in momentum, and v/t is the acceleration.

Force is then measured by *the rate of change of momentum*, or by *the product of the mass and the acceleration produced.*

Not only is the change of momentum the measure of the force producing it, but it always takes place in the direction in which the force acts. Both are vector quantities and their

directions are the same. The composition and resolution of forces are therefore effected in the same manner as those of other vector quantities (§§ 31, 32).

48. Units of Force. —Two systems of measuring force are in common use, the *gravitational* and the *absolute.* The latter is usually in the *c. g. s.* system. The gravitational unit of force is the *weight* of a standard mass, as the *pound of force,* or the *kilogram of force.* Gravitational units are not strictly constant, but vary with the place on the earth's surface (§ 59). They are not suitable, therefore, for precise scientific measurements.

The absolute unit of force in the *c. g. s.* system is the dyne. *The dyne is the force which produces an acceleration of one centimeter per second per second in a mass of one gram.*

The dyne is invariable in value. The earth's attraction for a gram mass in New York is approximately 980 dynes, since at that place gravity will impart to a gram an acceleration of 980 centimeters per second per second. A dyne is therefore $\frac{1}{980}$ of a gram of force, and the numerical value of any force expressed in dynes is 980 times as great as in grams of force (§ 52). Conversely, to convert dynes into grams of force, divide by the acceleration of gravity, 980. One kilogram of force is equivalent to 980,000 dynes.

49. Graphical Representation of a Force. —A force as a vector quantity has both *direction* and *magnitude ;* in addition it is often necessary to know its *point of application.* These three particulars may be represented by a straight line drawn through the point of application of the force, in the direction in which the force acts, and as many units in length as there are units of force. If a line 1 cm. long stands for a force of 1 dyne, a line 4 cm. long, in the direction *AB* (Fig. 19), will represent a force of 4 dynes acting in the direction from *A* toward *B.* Any point on the line *AB* may be used to indicate the point at which the force acts.

Fig. 19

Further, if it is desired to represent graphically the fact that the two forces act on a body at the same time, for example, one a force of 3 kgm. horizontally, and the other a force of 2 kgm. vertically, two lines are drawn from the point of application of the forces A (Fig. 20), one 3 units long toward the right, and the other 2 units long on the same scale toward the top of the page. The two lines, AB and AC, represent the two forces in point of application, direction, and magnitude.

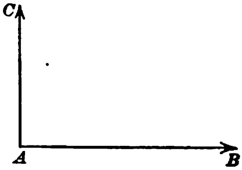


Fig. 20

50. How a Force is Measured. — A method of measuring a force, based on the relation $F = ma$, consists in measuring the mass moved by direct counterpoise (weighing) and observing the acceleration imparted to it by the force to be measured. But there are serious practical difficulties in measuring the acceleration.

The simplest method of measuring a force is by the use of an instrument known as a draw scale or spring balance, and in another form as a *dynamometer*. It consists essentially of a spring, to the free end of which is attached a pointer, arranged to move in front of a scale graduated in equal parts (Fig. 21). The form of the steel spring used is quite independent of the general principle that if two forces produce equal distortions of the spring, the forces themselves are equal to each other. If a weight of 20 pounds be hung on the spring and the position of the pointer be marked, then any other 20 pounds of force, whatever its origin, and in whatever direction applied, will stretch the spring to the same point. If a man by lifting stretches a spring two inches, and if a weight of 400 pounds stretches the spring to the same extent, then the man lifts with a force of 400 pounds of force.

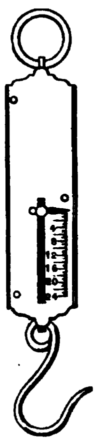


Fig. 21

The dynamometer, or spring balance, may be graduated in pounds of force, in kilograms or grams of force, or in dynes.

51. Discussion of the Third Law. — The essential meaning of the third law of motion is that all action between two bodies is *mutual.* It is a *stress*, and a stress always has two aspects, or is a two-sided phenomenon. The word *action* in Newton's third law is used to denote one aspect of a stress, and the word *reaction* the other.

Considered only with respect to one portion of a system of bodies a stress is called *action;* with respect to the remainder of the system it is called *reaction.* The third law states that these two phases of a stress are always equal and in opposite directions.

The stress in a stretched cord *pulls* the two things to which it is attached equally in opposite directions; the stress in a compressed rubber buffer exerts an equal *push* both ways. The former is a *tension;* the latter, a *pressure.*

The third law includes also what is sometimes called the *conservation of momentum*, and it may be expressed in modern phraseology as follows:

In every action between two bodies, the momentum gained by the one is equal to that lost by the other, or the momenta in opposite directions are the same.

When a bullet is fired from a gun, the momentum of the gun in one direction is equal to that of the bullet in the other. The velocities are not equal, but are inversely as the masses of the two.

52. Centripetal and Centrifugal Forces. — The acceleration in uniform circular motion has already been shown to be equal to v^2/r (§ 34). But force is the product of mass and acceleration ($F = ma$, § 47); hence *centripetal force* is equal to mv^2/r. It is the constant force in uniform circular motion that deflects a body from a rectilinear path, and compels it to move in a circle.

But by the third law of motion action and reaction are equal; hence the centripetal force has opposed to it the equal and opposite *centrifugal force.* The latter is the resistance which a body offers, *on account of its inertia,* to deflection from a straight path. We have then for either centripetal or centrifugal force the expression,

$$F = \frac{mv^2}{r}. \tag{16}$$

If m is in grams, v in centimeters per second, and r in centimeters, F is expressed in dynes. If it is desired to express F in gravitational units, divide the value obtained for F by the acceleration of gravity (§ 48). The result will be in grams of force or pounds of force, according to the units employed in equation (16).

To illustrate: If a mass of 500 gm. is made to revolve in a horizontal circle, whose radius is 1 m., and with a velocity of 3 m. per second, the centrifugal force is

$$F = \frac{500 \times 300^2}{100} = 45 \times 10^4 \text{ dynes; and } \frac{45 \times 10^4}{980} = 459.2 \text{ grams of force.}$$

Again, if a body having a mass of 5 pounds move in a circle of 10 feet radius with a velocity of 25 feet a second,

$$F = \frac{5 \times 25^2}{10 \times 32.15} = 9.72 \text{ pounds of force.}$$

(The acceleration of gravity in feet per second per second is 32.15.)

53. Illustrations of Centrifugal Force. — The stress, whose two aspects are centripetal and centrifugal force, exists in the medium or parts of the structure connecting the revolving member with the center. If a cord be just strong enough to sustain a weight of x gm. of matter, then it can sustain a stretching force of 980 x dynes. If this cord be used to whirl a mass of m gm. in a horizontal circle, it will snap unless the velocity is such that mv^2/r is less than 980 x dynes.

The stress in the cord consists on the one side of the centripetal force required to deflect the mass m from the tangent to the circle; on the other it is the centrifugal force of reaction which the body m exerts through the cord on the center.

Water adhering to the surface of a grindstone leaves the stone just as soon as the centripetal force, increasing with the velocity, is greater than the adhesion of the water.

An automobile rounding a curve at high speed is subject to strong centrifugal forces, which act through the tires. The centripetal force consists solely of friction between the tires and the ground. If this friction is insufficient, as on wet or icy ground, "skidding" ensues. In any case, rapid driving around curves brings to bear great lateral stresses on the tires.

Centrifugal machines are used in chemical laboratories to separate crystals from the mother liquors, in sugar refineries to separate sugar crystals from the syrup, and in dyeworks and laundries to dry yarn and cloth rapidly. Honey is extracted from the comb in a similar way. When light and heavy particles in a mixture are whirled, the heavier ones tend toward the outside, — are left behind in the rotation. Thus, the fat globules of milk, constituting the cream, are lighter than the liquid of the emulsion. Hence, when fresh milk is whirled in a dairy separator, the cream and the milk form distinct layers and are collected in separate chambers.

In Watt's steam engine governor, the balls open outward with increasing speed, thereby actuating a suitable train of mechanism, which throttles the steam at the inlet valve.

When a spherical vessel, containing some mercury and water, is rapidly whirled on its axis, both the mercury and the water rise and spread in separate bands as far from the axis of rotation as possible, the mercury outside.

The centrifugal force on a body may easily exceed its weight. If $mv^2/980\,r$ exceeds the weight of the body in grams, it will revolve in a vertical circle. Thus, a small open can, partly filled with water, may be whirled around in a vertical circle with an attached string without spilling the water. The "centrifugal railway" is similarly explained.

II. GRAVITATION

54. Free Fall of Bodies. — The erroneous notion of the early philosophers that heavy bodies fall faster than light ones was first corrected by Galileo. He dropped various bodies from the top of the leaning tower of Pisa, and found that they fell to the ground in nearly the same time, whatever their size or weight. The slight difference he observed he rightly ascribed to the resistance of the air.

Frictional air resistance is well illustrated by the "guinea and feather tube" (Fig. 22). If a small coin and either a feather or a pith ball are placed in the tube, the coin will fall to the bottom first when the tube is quickly inverted. But if the air is exhausted by a good air pump, the lighter object will fall as fast as the heavier one. In a perfect vacuum, all bodies at the same place on the earth's surface would show the same downward acceleration.

Fig. 22

The friction of the air against the surface of bodies moving through it leads to a limiting velocity. A cloud floats, not because it is lighter than the atmosphere, for it is actually heavier, but because the surface friction is so large in comparison with the weight of the minute drops of water, that the limiting velocity of fall is very small.

When a small stream of water flows over a high precipice, it is broken into fine spray and falls slowly. Such is the explanation of the Staubbach fall at Lauterbrunnen in Switzerland. The precipice is 300 m. high, and the fall viewed from the face resembles a magnificent transparent veil, kept in movement by currents of air.

In a vacuum water falls like a solid. The "water hammer" (Fig. 23) illustrates this fact. In filling the tube the water is boiled till all the air is expelled just before the tube is sealed in a blowpipe flame. When such a tube is suddenly inverted, the water falls like a solid and strikes the glass with a metallic ring. The same phenomenon may be observed when steam condenses in the cold pipes of a steam-heating system. The condensation produces a partial vacuum, and the water under steam pressure flows into it with a water hammer effect.

Fig. 23

55. Weight. — All the experimental evidence goes to show that every mass of matter at any given place on the earth's

surface will attain, when falling freely in a vacuum, the same velocity in a second. The force due to the earth's attraction, called *gravity*, is then proportional to the mass. The force with which the earth attracts a body is known in science as *weight*. The acceleration of gravity is denoted by g. Whatever may be the local value of the acceleration of gravity, the equation of force, $F = ma$ (§ 47) takes the form for gravity in the absolute system,

$$W = mg, \qquad (17)$$

where W is weight, m mass, and g the acceleration of gravity.

If the acceleration g varies from place to place, the local weight of a given mass varies in the same proportion.

56. Direction of Gravity. — The path described by a body falling freely is a *vertical line*. A line or plane perpendicular to it is said to be *horizontal*. The direction of the vertical at any point, which is nearly the direction in which gravity acts, may be determined by suspending a weight by a cord passing through the point. The cord suspending the weight is called a *plumb line*. The direction of the plumb line is perpendicular to the surface of still water.

A beautiful experimental demonstration of this fact consists in suspending a small weight by a thread so that the weight hangs under the surface of darkened water. The image of the thread in the water as a mirror may be distinctly seen, and it is exactly in line with the thread itself. But the image of a straight line in a plane mirror coincides in direction with the line itself only when the line is perpendicular to the mirror. Hence the thread as a plumb line is perpendicular to the surface of the darkened water.

Vertical lines drawn through neighboring points may be considered parallel without sensible error, for vertical lines 100 ft. apart make with each other an angle of only one second of arc; one second of arc is the angle subtended by a pinhead at the distance of about a quarter of a mile. At the poles of the earth and at the equator, the direction of gravity is that of the plumb line; elsewhere there is a slight deviation from this line on account of the rotation of the earth on its axis.

57. Newton's Law of Gravitation. — The famous astronomer Kepler discovered the laws of planetary motion, but he left untouched the forces which determine the motion. It remained for Sir Isaac Newton to give a dynamical explanation of Kepler's laws, and to show that a stress between each planet and the sun, directly proportional to the product of their masses and inversely proportional to the square of the distance between them, would account for the planetary motions according to Kepler's laws. The law of universal gravitation enunciated by Newton is:

Every portion of matter attracts every other portion, and the stress between them is proportional to the product of their masses and inversely proportional to the square of the distance between them.

The law expressed in symbols is

$\frac{L^3}{MT^2}$

$$F = G\frac{mm'}{d^2}, \tag{18}$$

where m and m' are the two attracting masses, d the distance between them, and G a proportionality factor or the constant of gravitation. For spherical bodies the distance d in the law of gravitation is the distance between their centers. It is readily shown* that the attraction at any external point due to a sphere, either uniform in density (§ 154) or made up of concentric layers each uniform throughout, is the same as if the mass of the sphere were collected at its center. By the attraction at a point is meant the attraction on unit mass at the point.

58. Law of Gravitation applied to the Moon. — Given the acceleration of gravity at the surface of the earth, the law of inverse squares gives the acceleration produced by the gravitational attraction of the earth at the distance of the moon. Let g' be this latter value, and let R be the distance of the moon, and r the radius of the earth; then

$$g : g' = R^2 : r^2.$$

Whence

$$g' = g\frac{r^2}{R^2}.$$

* *University Physics*, Part II, Art. 121.

If the law of gravitation is true, g' should be equal to the centripetal acceleration of the moon in its orbit, for the attraction of the earth for the moon is the only force present to keep the moon in its orbit.

From equation (12) the centripetal acceleration of the moon in its orbit is given by the equation

$$a = \frac{4\pi^2 R}{T^2},$$

where T is the period of the moon, or the lunar month. Then if the law of gravitation is true, g' should be equal to a.

The following are the necessary data for the approximate calculation of g' and a:

$$R = 240{,}000 \text{ miles},$$
$$r = 4000 \text{ miles},$$
$$T = 27 \text{ da. } 8 \text{ hr.} = 2{,}361{,}600 \text{ sec.},$$
$$g = 32.2 \text{ ft. per second per second.}$$

Then

$$g' = 32.2 \frac{(4000)^2}{(240{,}000)^2} = 0.00894 \text{ ft. per second per second.}$$

Also $$a = \frac{4\pi^2 \times 240{,}000 \times 5280}{(2{,}361{,}600)^2} = 0.00898 \text{ ft. per second per second.}$$

The agreement is good for the approximate data used, and it shows that the moon is attracted by the earth with a force which follows the law of universal gravitation.

59. Law of Weight.—Since the earth is flattened at the poles, it follows that the acceleration of gravity, and the weight of any body, increase in going from the equator toward the poles. If the earth were a uniform sphere and stationary, the value of g would be the same all over its surface. But the value of g varies from point to point on the earth's surface, even at sea level, both because the earth is not a sphere and because it rotates on its axis. The value of g at the equator is 978.1 and at the poles 983.1, both in centimeters per second per second. At New York it is a little over 980 cm. per second per second, or 32.15 ft. per second per second.

The diminution of gravity at the equator on account of the rotation of the earth on its axis is easily calculated. The equatorial radius of

the earth is 6378×10^5 cm. The period of rotation is a sidereal day of 86,164 sec. Hence

$$a = \frac{4\pi^2 \times 6378 \times 10^5}{(86,164)^2} = 3.4 \text{ cm. per second per second.}$$

$$3.4 = \frac{1}{289}g = \frac{1}{17^2}g.$$

Since the centripetal acceleration varies inversely as the square of the period of revolution, it would equal g at the equator if the period of rotation of the earth were reduced to one seventeenth of a day. The apparent acceleration of gravity at the equator would then be reduced to zero, or bodies there would have no weight.

60. Laws of Falling Bodies. — The most familiar and important example of uniformly accelerated motion, the formulae for which have already been given in § 29, are presented by falling bodies. Since the acceleration g is sensibly constant for small distances above the earth's surface, the formulae of § 29 are directly applicable by substituting for a the specific acceleration g. Equations (7), (8), and (9) then become

$$v = gt, \tag{19}$$

$$s = \tfrac{1}{2} gt^2, \tag{20}$$

$$v^2 = 2\,gs. \tag{21}$$

If in (20) t is made one second, $s = \frac{1}{2}g$; or the space described in the first second, when the body starts from rest, is half the value of acceleration of gravity. A body falls 490 cm. the first second; the velocity attained is 980 cm. per second, and the acceleration is 980 cm. per second per second.

The following laws are embodied in the above equations:

I. *The velocity attained by a falling body is proportional to the time of falling.*

II. *The space described is proportional to the square of the time.*

III. *The acceleration is twice the space through which a body falls in the first second.*

61. Projection Upward. — When a body is thrown vertically upward, the acceleration is negative, and its velocity decreases each second by g units (980 cm. or 32.15 ft.). The time of ascent to the highest point will be the time taken to bring the body to rest. If the velocity lost is g units per second, the time required to lose v units of velocity will be the quotient of v divided by g, or

$$t = \frac{v}{g}. \tag{22}$$

If, for example, the velocity of projection vertically upward is 1960 cm. a second, the time of ascent, neglecting atmospheric resistance, is $\frac{1960}{980}$, or 2 seconds. This is the the same as the time of descent again to the starting point.

62. Center of Gravity. — A body is conceived to be composed of an indefinitely large number of parts, each of which is acted on by gravity. For bodies of ordinary size these forces of gravity are parallel and proportional to the masses of the several small parts. The point of application of their resultant is the *center of gravity* of the body. This point is also called the *center of mass* and the *center of inertia.*

If the body is of uniform density throughout, the position of its center of gravity depends on its geometrical figure only. Thus, the center of gravity (1) of a straight line is its middle point; (2) of a circle or ring, its center; (3) of a sphere or a spherical shell, its center; (4) of a parallelogram, the intersection of its diagonals; (5) of a cylinder, the middle point of its axis.

It is necessary to guard against the idea that the force of gravity on a body acts at its center of gravity. Gravity acts on all the constituent particles of a body; but its effect is generally the same as if the resultant, that is, the entire weight of the body, acted at its center of gravity.

63. Center of Gravity of a Triangle. — A single example will serve to illustrate the geometrical method of finding the center of gravity of a figure uniform throughout.

Let ABC (Fig. 24) be a uniform thin triangle. Draw a line from A to D, the middle point of the base. The triangle may be conceived to be made up of lines parallel to its base BC; and since the bisector AD passes through the middle point of all these lines, the center of gravity of each line lies on AD, and therefore the center of gravity of the whole triangle also lies on it.

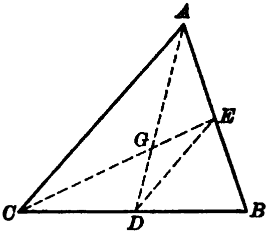


Fig. 24

Let CE be another bisector drawn from C to the middle point of the side opposite. For reasons similar to those above, the center of gravity of the triangle must lie also on the line CE. It must therefore be at the intersection of the two bisectors.

The triangles DEG and ACG are similar. But DE is one half of AC; therefore DG is one half of AG, or one third of AD. The center of gravity of a triangle is therefore on a line drawn from the apex to the middle point of the base and two thirds of the distance down.

III. WORK AND ENERGY

64. Work Defined. — When a force acts on a body and produces displacement in the direction in which it acts, the force is said to do *mechanical work*.

Examples: Gravity does work on the weight of a pile driver, causing it to descend; steam exerts pressure on the piston of a steam engine, imparts motion against resistance, and does work; a horse does work in pulling a wagon up an inclined roadway; the electric current, by means of a motor, does work when it drives an air compressor and forces air into a compression tank.

Unless the point of application of the force has a component motion in the direction of the force, no work is done. Thus, gravity does no work on a vessel moving over the level surface of the sea, because its motion is at right angles to the direction of the force of gravity; the pillars support-

ing a pediment over a portico do no work, though they support a weight and exert a force. The forces are balanced and there is no motion.

The measure of mechanical work is the product of the force and displacement of its point of application in the direction of the force, or

$$W = Fs. \tag{23}$$

Since force is the product of mass and acceleration (§ 47),

$$W = mas. \tag{24}$$

When the displacement produced makes an angle α with the direction of the force,

$$W = Fs \cos \alpha. \tag{25}$$

This expression may be interpreted to mean either, (a) the product of the force (F) and the component of the displacement in the direction of the force ($s \cos \alpha$), or (b) the product of the displacement (s) and the component of the force in the direction of the displacement ($F \cos \alpha$).

65. Units of Work.—Three units of work are in common use:

1. The *foot pound*, or the work done by a pound of force working through a distance of one foot. This unit is in general use among English-speaking engineers.

2. The *kilogram meter*, or the work done by a kilogram of force working through a distance of one meter. This is the gravitational unit of work in the metric system.

3. The *erg*, or the work done by a force of one dyne working through a distance of one centimeter. The erg is the absolute unit in the *c. g. s.* system and is invariable.

Gravity gives to the gram a velocity of about 980 cm. a second. It is therefore equal to 980 dynes; and if a gram mass be lifted vertically one centimeter, the work done against gravity is one gram centimeter, or 980 ergs.

The mass of a "nickel" is 5 gm. The work done in lifting it vertically two meters is the continued product of 5, 200,

and 980, or 980,000 ergs. The erg is therefore a very small unit, and it is more convenient to use a multiple for practical measurements. The multiple in common use is the *joule*. Its value is

$$1 \ joule = 10^7 \ ergs = 10,000,000 \ ergs.$$

Expressed in this larger unit, the work done in lifting the "nickel" is 0.098 joule.

66. Graphical Representation of Work.—Since work is the product of force and length, work may be represented numerically by an area. When the force is constant in value, the work done may be denoted by the area of a rectangle, one side of which is as many units in length as there are units of force, while the adjacent side is numerically equal to the displacement in the direction of the force (Fig. 25).

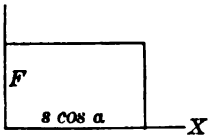


Fig. 25

If the force increases uniformly from zero to a final value F, then the work done is the product of the mean value of the force and the displacement. It may be represented by the area of a right triangle (Fig. 26), in which the base is the displacement in the direction in which the force acts, and the altitude the final value of F; for the work then equals $\frac{1}{2} Fs \cos \alpha$, which is the expression for the area of the triangle.

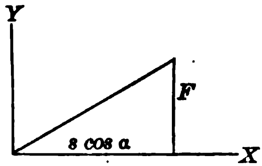


Fig. 26

In many cases, as in the cylinder of a steam engine, the force, which is now the pressure of the steam, varies according to some law less simple. If p is the pressure per unit area of the piston, and A its area, then the whole pressure on the piston is $P = pA$. Let now the piston move through a very small distance x, so small that the pressure may be considered constant during this motion; then the work done by the steam during expansion is

$$w = pAx.$$

But Ax is the increase in the volume of the steam, which may be denoted by v. Then

$$w = pv,$$

or the element of work done during a very small movement of the piston is equal to the product of the pressure per unit area and the small change in volume.

The whole work done during any considerable expansion of the steam may be represented by the area of the figure $ABba$ (Fig. 27), in which the ordinates of the curve AB are the successive pressures, and the abscissas are the corresponding volumes of the steam. Take any small element of the area, as Aa. Then the length of this strip is the instantaneous pressure p, and its width is the indefinitely small change in volume. Its area is therefore the element w of work done, and the sum of all such elements is the entire work done by the steam during the expansion from the volume denoted by the point b to that denoted by the point a. The sum is the area $ABba$. This principle is the one employed in the steam indicator diagram.

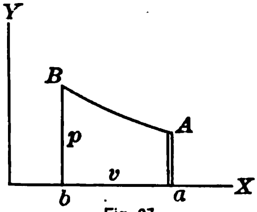


Fig. 27

67. Power. — *Power is the time rate of doing work.* The unit of power commonly used by American and British engineers is the *horse power;* it is the rate of doing work equal to 33,000 foot pounds per minute, or 550 foot pounds per second.

In the *c. g. s.* system the unit of power is the *watt.* It is the work done at the rate of one joule (10^7 ergs) per second. A *kilowatt* is 1000 watts.

To convert a horse power into watts, multiply 550 by the ratio between a foot and a centimeter, then by the ratio between a pound and a gram, and the product is in gram centimeters. To find the equivalent in ergs multiply by 980. Then $550 \times 30.4797 \times 453.59 \times 980 = 746 \times 10^7$ ergs per second, or 746 watts.

One horse power is therefore equivalent to 746 watts. A kilowatt (K. W.) is nearly one and one third horse powers.

To convert kilowatts into horse powers, add one-third; to convert horse powers into kilowatts, subtract one fourth. For example, 90 K. W. equals 120 H. P., and 200 H. P. equals 150 K. W.

The capacities of electric generators are now universally expressed in kilowatts; electric motors and steam engines are also commonly rated in the terms of the same unit of power.

The horse power was originally determined by James Watt, and the average work per horse turned out to be 22,000 foot pounds per minute. For some reason Watt added fifty per cent and made the horse power 33,000 foot pounds per minute. (Report of the Electrical Conference, Philadelphia, 1884.) If he had adopted the value found of 22,000 foot pounds per minute, one horse power would have been almost precisely one half kilowatt.

68. Energy. — Experience teaches that under certain well-defined circumstances bodies possess the capacity for doing work. Thus a body of water in an elevated position, air under pressure in a tank, steam under more than atmospheric pressure in a steam boiler, are all able to do work by means of an appropriate motor mechanism. In general, a body or system upon which work has been done is found to have an increased capacity for doing work. It is then said to possess more *energy* than before. *Energy is the capacity for doing work.* It is therefore measured in the same units as work.

When a steam engine lifts the weight of a pile driver, it does work on it against the force of gravity. In its new position relative to the earth the weight itself has acquired the ability to do work; for when it is released, it descends, overcomes the resistance offered by the pile, and forces it into the ground.

Work may be done on a storage battery by means of a steam engine and a dynamo machine. The charged battery has then conferred on it the capacity for doing work, because it is capable of furnishing an electric current to run a motor. It may circulate the air by a fan, drive a printing press, run a street car, propel an electric launch, or operate the machinery of a factory.

Consider examples of a different character. Work is done on a cannon ball by means of the pressure of the gases arising from the explosion of the powder. The ball acquires a high speed; and, as a result, it now

possesses the capacity of overcoming resistance. By virtue of its mass and its motion it may demolish fortifications, or pierce the steel plates of a battleship.

When steam does work on the piston of an engine, the heavy flywheel is made to revolve on its axis. Work is done on it in giving it motion of rotation. When the steam is shut off, the engine continues to revolve and does work by means of the rotatory effort of the massive flywheel.

In all these cases while the acting agent is doing work on the body, *energy* is transferred from it to the body or system on which the work is done; and the body or system of bodies which has acquired the capacity for doing work is said to possess *energy*.

69. Potential Energy. — Cases abound in which energy is stored in mechanical displacements, or in chemical and physical changes in a body. The first two illustrations of the last section belong to this class. In the air gun, a mass of air is compressed into a small volume by doing work on it, and it tends to recover its original dimensions; if permitted to do so, it may be made to restore the work done on it by propelling a bullet. When a clock is wound by coiling a spring or lifting a weight, work is done on the system and energy is stored. This energy is recovered when the system returns slowly to its unstrained condition by the uncoiling of the spring, or the fall of the weight. The work done in bending a bow is quickly restored in imparting motion to the arrow as the bow is relieved from the stress.

In all such cases of the storage of energy a stress is always present. The compressed air pushes outward in the air gun; the spring strives to uncoil in the clock; the bent bow tends to unbend; and the electric pressure of the charged storage battery is ready to produce a current as soon as the circuit is closed. Hence the energy thus acquired is called *energy of stress*, or more commonly *potential energy*. The energy of an elevated mass, of bending, twisting, deformation, of chemical separation, and of stress in the ether in a magnetic field are all cases of potential energy.

70. Kinetic Energy.—Bodies have capacity for doing work also in consequence of their motion. In the last two illus-

trations of § 68, the immediate and obvious effect of doing work on the body is to set it moving, but in reality energy is imparted to it. The energy which it acquires is called *kinetic energy*, or energy of motion.

Whenever a meteoric body, flying through space, enters the earth's atmosphere, its energy of motion is converted into heat by friction with the air, and the heat generated raises its temperature till it glows like a star. It may even burn up or become impalpable powder.

The invisible molecular motions of bodies constituting heat are included under kinetic energy no less than their visible motions. Heat is a form of kinetic energy.

Kinetic energy must not be confused with force. A moving mass of matter carries with it a definite quantity of energy, but it exerts no force until it encounters resistance or opposition. Energy is then transferred to the resisting or opposing body; force is exerted only during this transfer.

71. Kinetic Energy in Terms of Mass and Velocity. — Suppose a force F to act on a body of mass m for an interval of time t; then the measure of the effect is the impulse Ft. By the second law of motion impulse equals the momentum imparted. Assuming that the body m starts from rest and acquires in time t a velocity v, the momentum produced is mv. Hence

$$Ft = mv. \tag{a}$$

A constant force gives rise to uniformly accelerated motion, and in this type of motion the mean velocity is half the sum of the initial and final velocities, or, in this case, $\frac{1}{2}v$. The mean velocity is also the space traversed divided by the time, or $\frac{s}{t}$. Hence

$$\frac{s}{t} = \frac{1}{2}v. \tag{b}$$

Multiply (a) and (b) together, member by member, and the result is

$$Fs = \frac{1}{2}mv^2. \tag{26}$$

But Fs measures the work done by the force F on the mass m to give to it the velocity v, while working through the

space s; and as the kinetic energy acquired by the body is measured by the work done on it, it follows that the energy of the mass m moving with the velocity v is $\frac{1}{2}mv^2$.

If m is expressed in grams and v in centimeters per second, the result is in ergs. To reduce to gram centimeters, divide by the value of g in centimeters per second per second, 980. If m is expressed in pounds and v in feet per second, to obtain the energy in foot pounds, divide by the value of g in feet per second per second, 32.15.

72. Energy changes Form. — The energy of a body may change from potential to kinetic, and conversely. Suppose work has been done on a weight of mass m gm. sufficient to lift it to a height of h cm. against gravity. It then possesses potential energy equal to mgh ergs. If it is allowed to fall, it loses potential energy and gains energy of motion. After it has fallen a distance s, its velocity is given by the equation $v^2 = 2gs$. Its kinetic energy is then $\frac{1}{2}mv^2 = mgs$. But its potential energy has been reduced to $mg(h - s)$, since $h - s$ is now its height above the point from which it was lifted. The sum of mgs and $mg(h - s)$ is mgh, the original potential energy. Whatever, then, the weight gains in kinetic energy as it falls, it loses in energy of the potential form. When the weight reaches the ground, the velocity acquired is given by the relation $v^2 = 2gh$, and therefore $\frac{1}{2}mv^2 = mgh$, or the energy of the weight is now all kinetic and is the same as the potential energy possessed by the weight at the elevation h. During the fall the potential energy is continuously converted into the kinetic form, but in such a way that the sum of the two is a constant.

When a pendulum is drawn aside, it is lifted and acquires potential energy. As it descends toward its position of equilibrium it acquires velocity, and at the lowest point of its path, its energy is all energy of motion. Its energy then gradually returns to the potential form, and at the extremity of the swing on the other side it is again all potential. If the pendulum could swing without friction and resistance of

the air, this process of conversion of energy from the one form into the other and back again would continue indefinitely without loss. This would constitute a form of perpetual motion, but it would be one in which no energy is given out to other bodies. It is not the form sought after by the deluded.

These forms of potential and kinetic energy of a mass of matter, such as a pendulum bob, are not the only ones assumed by energy. When the lifted weight falls and reaches the ground, its motion may be suddenly arrested, and its energy of visible mechanical motion disappears. What becomes of it? It is found that both the weight and the ground are warmed, and by a wide induction from similar cases, it is learned that heat is a form of energy. The same quantity of work, if entirely spent in producing heat, will always produce the same amount.

When a ball strikes a target, heat is generated, and the ball may be partly fused. A flash of light is often noted and sound is produced. The molecular motions constituting heat, light, and sound represent kinetic energy.

The heat and light produced by combustion are forms of kinetic energy, derived by transformation from the potential energy of chemical separation and chemical affinity.

When a storage battery is charged, the energy is stored as the potential energy of chemical separation. When it discharges, the kinetic energy of the current flowing is derived by transformation from the potential form.

73. Conservation of Energy. — The principle of the *conservation of energy* is a generalization from an extensive range of observations. It has been found that if a system of bodies within a given boundary, through which energy is not allowed to pass, has a certain amount of energy, this amount remains constant, whatever actions take place between the parts of the system, and whatever forms this energy may assume.

If this system is brought into relation with other bodies, so as to form a larger system, in which the energy is differently distributed, then the entire energy of the two systems within the larger boundary remains invariable in amount.

So if the boundary is extended to include the whole physical universe, the doctrine of the conservation of energy amounts to the assertion that the quantity of energy in the universe is fixed and invariable.

It will be seen that this principle denies the possibility of any form of "perpetual motion" machine by which mechanical work can be done continuously without supplying the machine with equivalent energy in some other form. Every machine or device that works without interruption must receive from without at least as much energy as it expends, either continuously or by periodic additions.

74. Availability of Energy. — While the quantity of energy in the universe remains unaltered, its *availability* for the operations of nature and for purposes useful to man is not a constant.

Whenever a transformation of energy occurs, especially from the kinetic to the potential form, some of it is inevitably wasted through conversion into heat. This conversion, or really degradation of energy, occurs through friction, radiation, the heating of electric conductors, or other analogous modes. This heat is gradually diffused, and diffused heat is no longer available as useful energy to do work.

Not only this, but all the processes of nature exhibit changes of energy on the way from the more available to the less available form. Hence the quantity of available energy is never increased, but is always diminished in every physical process, and it therefore tends toward zero.

The constant dissipation of energy associated with every physical phenomenon leads to two interesting conclusions. In the first place, if we permit ouselves to inquire into the past history of energy, we shall inevitably arrive at a period in the past when none of it had been dissipated or had become unavailable. Before this period no physical phenomenon, like those with which we are acquainted, could have occurred, for every such phenomenon is attended with dissipation of energy.

In the second place, unless some new order intervenes, of which we have no conception, we are forced to contemplate a moment in the

distant future when all energy will be in the unavailable form of equally diffused heat, and the whole physical universe will have run down like the weights of a clock.

IV. FRICTION

75. Kinds of Friction. — A body in motion relative to another body, with which it is in contact, or to a medium through which it moves, is always subject to retarding forces tending to bring it to rest. This action is called *friction.* It may take several forms:

1. *Sliding friction.* This occurs when the surface of one body slides along that of another. Sliding friction opposes the motion of many parts of machines, such as the sliding motion of the crosshead of a steam engine between its guides, or its shaft sliding round and round in its bearings.

2. *Rolling friction.* Whenever a wheel or a cylinder rolls on a plane surface, the resistance to motion at the line of contact is called *rolling friction.* Rolling friction is probably due to the yielding of the wheel and the surface on which it rolls. The effect is much the same as if the wheel were continually climbing a slight ascent.

3. *Fluid friction.* The air offers resistance to the passage of a body through it; a pipe offers resistance to the flow of steam or water through it; water offers resistance to the motion of a boat, or the flight of a torpedo through it. These resistances are called *fluid friction.*

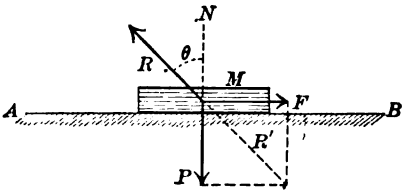


Fig 28

76. Coefficient of Sliding Friction. — Suppose a body M (Fig. 28) on a horizontal plane AB. Let a force F act on it parallel to the face of the plane. Also let P be the force vertically downward, pressing the surfaces of M and AB together. P may be the weight of the body M, or it may have any other value or source. The resultant of F and P is R'; and

if M remains at rest, the reaction of the plane AB must be equal and opposite to R', making an angle θ with the normal N.

The body M will not begin to slide unless F is a certain fraction of P. This fraction is called the *coefficient of statical friction.* It is usually denoted by μ. Then

$$F = \mu P. \tag{27}$$

The coefficient μ has to be determined experimentally.

The force F encounters a frictional resistance which has a maxium value μP. If F is less than μP, sliding does not occur; but as F is increased, the frictional resistance also increases up to the value μP, when slipping begins. It cannot increase beyond this limiting value, and is slightly reduced after the body begins to slip.

The coefficient μ depends on the nature of the surfaces in contact, on their condition as to smoothness or roughness, and on the presence or absence of lubricants. It is independent of the area of the surfaces in contact, and of the pressure P. The whole frictional resistance is therefore proportional to the pressure P. But as this remains constant while the area of contact is diminished, the frictional resistance per unit area is directly as the pressure.

77. The Limiting Angle. — If F in Figure 28 is the value of the frictional resistance just as slipping begins, then the angle θ between R and N is the *limiting angle* of friction. It is the angle at which the reaction of the plane is inclined to the vertical. Since the horizontal component of R' is F and its vertical component P, we have $\tan\theta = \frac{F}{P} = \mu$; that is, the tangent of the limiting angle is equal to the coefficient of friction. If there were no friction, any force applied to M, deviating ever so little from the vertical, would produce slipping; but since there is friction, to produce sliding the force applied must make an angle with the normal to the surface of contact

somewhat larger than the limiting angle. Any pressure, however great, so long as its direction lies within the cone described by rotating the line R about the normal N as an axis, is incapable of making the body slip.

Friction continues after the sliding motion has begun and opposes motion; but its magnitude is somewhat less than the friction of rest at the moment the slipping begins.

78. **Angle of Repose.** — Let a body M (Fig. 29) be placed on a plane surface, and let this surface be tilted until the body is on the point of beginning to slide. Let θ be the angle of elevation of the plane. Then the weight of M equal to mg acts vertically downward, and the friction F is directed upward parallel to the face of the plane.

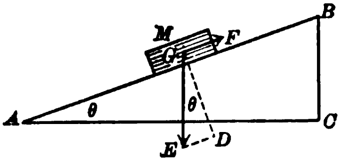


Fig. 29

Resolve mg into a normal component GD, which is offset by the reaction of the plane, and a component DE, which tends to make the body M slide down the plane. The friction F must exactly balance this latter component at the instant when slipping begins. The angle θ is then called the *angle of repose*. It is the same as the limiting angle. For the pressure P is the component GD equal to $mg \cos\theta$. The component DE, representing F, is equal to $mg \sin\theta$. Hence

$$\frac{mg \sin\theta}{mg \cos\theta} = \frac{F}{P} = \tan\theta = \mu,$$

or

$$F = \mu P.$$

The angle of repose is the angle at which heaps of sand, grain, or even fruit and shot, adjust themselves when dumped and allowed to find their own position of rest under sliding friction.

79. **Rolling Friction.** — When a cylindrical body rolls over a surface, the frictional resistance to motion is less than for

sliding. This fact explains the conspicuous advantage of wheeled vehicles over sledges. Rolling friction is equivalent to a small force acting at the circumference of the cylinder and bearing a small ratio to the pressure of the cylinder on the surface.

This ratio is affected by the relative yielding of the surfaces in contact. Osborne Reynolds has shown that an iron wheel rolling on india rubber raises before it a little hummock. This hummock tends to recover its form and to drive the wheel backwards. The rolling friction of iron on india rubber was found to be ten times as great as of iron on iron.

Conversely, a rubber tire (automobile) is visibly deformed when rolling on a hard surface. The part in contact with the hard surface changes length, and hence there is some slip between the two. This action means a thrust on the contact layer of both substances. In the case of iron rails, it results in the constant peeling off of thin scales of iron.

When the motive power is applied to rotate the wheel, as in the drivers of a locomotive or the rear wheels of an automobile, both rolling and slipping may be present. The friction between the driving wheels of a locomotive and the rails acts *forward*, while the rail itself is pulled backward on the ties and ground. This friction constitutes the external force pulling the train forward. The bottom of the drivers is forced backward by the pressure of the steam on the pistons, and friction on the rails resists this motion. On the other hand, the railway coaches are pulled forward; friction opposes this motion and causes the bottom of the wheels to turn backward. The friction between the drivers and the rails is therefore necessary to motion forward. Without it, the wheels of the locomotive would simply spin around without moving the train. This is what sometimes happens when the rails are wet or covered with ice.

Fig 30

If the brakes are set too hard, the friction between the carriage wheels and the rails may be insufficient to turn them against the brake friction. The wheels are then set and the breaking action is as great as possible. In this way a railway carriage may actually slide down a heavy grade without a wheel turning.

The friction of a round solid rolling on a smooth surface is always less than when it slides. Advantage is taken of this fact to reduce the friction of bearings. A ball bearing (Fig. 30) substitutes the rolling friction between balls and rings for the sliding friction between a shaft and its journal.

80. Loss of Energy Due to Friction. — Friction acts in general as a resistance opposing motion. Whenever a displacement takes place against frictional resistance work must be done. The energy equivalent to this work is converted into heat, which is gradually diffused among neighboring bodies, and the energy so transformed is no longer available to do work. Friction, therefore, decreases the efficiency of machinery by wasting energy.

81. Friction Dynamometer. — An extreme case of the absorption of energy by friction is presented by the *friction dynamometer*. It is a device for measuring the power of a steam or a gas engine, or of an electric motor.

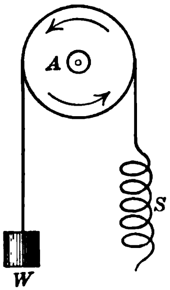


Fig. 31

In Figure 31, A is a pulley fixed to the revolving shaft of the engine or motor. It has a channeled rim, the bottom of the channel being flat. Around it is wrapped a cord with one end attached to a spring balance S, which measures the tension in that end of the cord. A tension represented by the weight W suspended on the other end keeps the cord straight and determines the load. When the pulley is running, the work done is all absorbed by the friction between the stretched cord and pulley.

Let r be the radius of the circle at the axis of the cord. The force of friction worked against is the tension F, indicated by the spring balance S, less the weight W, since this weight acts in the same direction as the pulley turns. The distance worked through in one revolution is $2\pi r$. If the speed is n revolutions per second, the distance worked through

by the pulley in a second is $2\pi nr$. The work done per second, or the power, is then

$$P = (F - W) 2\pi nr.$$

If F and W are in pounds and r is in feet, the horse power is obtained by dividing by 550. If F and W are in dynes (gm. $\times$ 980) and r is expressed in centimeters, the result will be found in watts by dividing by 10^7, or in kilowatts by dividing by 10^{10}.

The quantity $2\pi n$ is the angular velocity ω (§ 87), because it is the angle in radians described per second. $(F - W)r$ is called the *torque*, T. The power is therefore equal to $T\omega$, or the product of the torque and angular velocity.

V. MOTION OF ROTATION

82. Rotation about a Fixed Axis. — An unbalanced force applied to a rigid body will in general produce both motion of translation and motion of rotation, unless it is directed through the body's center of mass, when the motion will be one of pure translation (§ 25). In order, then, to study motion of rotation by itself, we may assume that the body has in it a fixed line, so that the only motion possible for it is rotation about this line as an axis. Under these conditions a single unbalanced force produces rotation only. An example is the flywheel of an engine, or a door swinging on its hinges. All points in the body then describe circles about the fixed axis as a center, and all have the same *angular* velocity.

83. Moment of a Force. — The effect of a force F in producing rotation is dependent not only on the magnitude of the force, but also on the distance of its line of action from the axis of rotation. It is obvious that a smaller force is required to close a door when applied at right angles to the door at the knob, than when applied near the hinge. Also, an increase in the rotatory effect on a flywheel may be secured either by increasing the force, or by lengthening the crank.

The effectiveness of a force in producing rotation depends on two quantities: (*a*) the magnitude of the force, and (*b*) the shortest distance between the axis and the line of action of the force.

The measure of this effectiveness is the product of the force and the perpendicular distance between its line of action and the axis of rotation. This product is called the *moment of the force*. When the force is applied to turn a shaft, the moment is usually called the *torque*.

Let M (Fig. 32) be a rigid body which may rotate about an axis through O perpendicular to the plane of the figure. Then the moment of the force F is not so great if it is applied at B in the direction AB, as if the same force is applied at B in the direction CB. The moment for the first direction is $F \times OA$, and for the second, $F \times OB$.

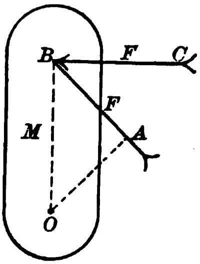


Fig. 32

A moment is considered positive if it tends to produce rotation in one direction, and negative if in the other. It is immaterial which direction is considered positive if the same direction remains positive throughout any one problem.

84. The Moment of the Resultant equals the Algebraic Sum of the Moments of the Components. — If the resultant of any number of forces is a single force producing the same effect as the component forces acting conjointly, then the moment of this resultant about any point should equal the algebraic sum of the moments of the several components about the same point; otherwise the resultant could not replace the components in producing rotation. (Note the special case of a couple, § 86.)

The principle applied to a parallelogram of forces may be demonstrated as follows: Let two forces P and Q be represented by the lines AP and AQ, and their resultant R by the

diagonal AR of the parallelogram (Fig. 33). Let O be any point in the plane of the figure as the point about which the moments are to be taken. It is called the *center of moments.*

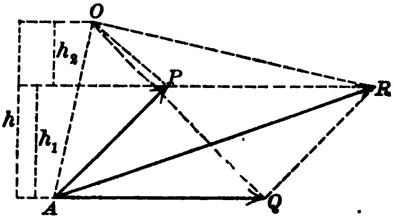


Fig. 33

The area of the triangle AOQ is half the moment of the force Q; for the moment of Q is the product of the lines AQ and h, while the area of the triangle is half this product. In the same way it may be shown that the moment of P is twice the area of AOP, and the moment of R is twice the area of AOR. We are to prove then that the area AOR equals the sum of the areas AOP and AOQ.

$$\text{Area } AOQ = \text{Area } APR + \text{Area } OPR,$$

since the three triangles have equal bases AQ and PR, and the altitude h of the first is equal to the sum of the altitudes h_1 and h_2 of the other two. Also, since the whole is equal to the sum of its parts,

$$\text{Area } AOR = \text{Area } AOP + \text{Area } APR + \text{Area } OPR.$$

Hence, substituting for the last two areas their equivalent area AOQ,

$$\text{Area } AOR = \text{Area } AOP + \text{Area } AOQ.$$

Therefore the moment of the resultant R is equal to the sum of the moments of the components P and Q.

If the point O, about which the moments are taken, lies on the line denoting the direction of the resultant, the moment of the resultant for that point is zero, and the algebraic sum of the moments of all the component forces is zero; or the sum of the positive moments in one direction is equal to the sum of the negative moments in the other.

85. Parallel Forces. — An important case for the application of moments is that of parallel forces.

The resultant of several parallel forces is equal to their algebraic sum and acts in a direction parallel to them (§ 31).

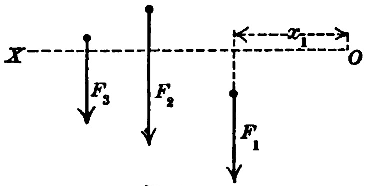


Fig. 34

Let F_1, F_2, F_3, etc., be any number of parallel forces (Fig. 34), and OX any line perpendicular to their directions. Also let x_1, x_2, x_3, etc., be the distances from any point O on OX to the forces respectively. Then by the principle of moments

$$RX = F_1x_1 + F_2x_2 + F_3x_3 + \dots,$$

or
$$X = \frac{F_1x_1 + F_2x_2 + F_3x_3 + \cdots}{F_1 + F_2 + F_3 + \cdots} = \frac{\Sigma Fx}{\Sigma F}. \qquad (28)$$

R is the resultant and X is its distance from O. (The sign of summation Σ is read "the sum of such terms as.")

Specifically, if there are only two forces, F_1 and F_2, acting in the same direction, let the point O be on the resultant itself. Then X is zero and
$$0 = F_1x_1 - F_2x_2,$$

or
$$\frac{F_1}{F_2} = \frac{x_2}{x_1}.$$

Now since $x_1 + x_2$ is the distance d between the parallel forces, it is obvious that the resultant divides this line into parts inversely as the forces.

If the parallel forces are in opposite directions, and F_1 is the greater, the resultant is $F_1 - F_2$. Take the point O on F_2. Then
$$X = \frac{F_1d}{F_1 - F_2}.$$

But $F_1 - F_2$ is necessarily less than F_1, and X is therefore greater than d; *i.e.* the resultant lies outside the forces and on the side of the greater.

Parallel forces may be illustrated by means of a graduated bar balanced on a knife-edge at O (Fig. 35). The weight of the bar itself does not

produce rotation if its center of gravity coincides with the knife-edge. A weight placed on the right side produces a rotation clockwise; one on the left, rotation counter-clockwise.

The two weights, W_1 and W_2, placed at distances 5 and 8 units respectively from O, will balance each other if their moments about O are the same, *i.e.* if

$$W_1 \times 5 = W_2 \times 8,$$

or if $$W_2 = 5/8 \times W_1.$$

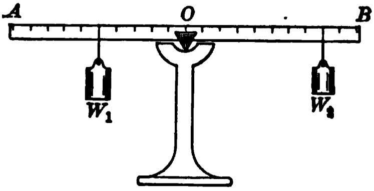


Fig. 35

Also, any number of weights on one side will balance any number on the other if the sum of the moments of those on one side is equal to the sum of the moments of those on the other. Whenever the bar is balanced the resultant of all the weights passes through the knife-edge; hence the moment of the resultant about this point is zero, and the sum of all the moments in one direction must equal the sum of all those in the other.

86. Couples. — Two equal parallel forces acting in opposite directions constitute *a couple*. The resultant of a couple is zero. It does not produce motion of translation nor acceleration of the center of mass.

A couple is simply a rotator, and its moment is equal to the product of one of the forces and the perpendicular distance between them, Fd. The perpendicular distance d is called the *arm* of the couple. The moment of the couple is therefore the same about any point in its plane.

All couples in the same plane are equivalent to one another if their moments are equal, because they produce the same rotational effect about any point in their plane. A couple may then be supposed moved to any part of its plane.

The resultant of any number of couples in one plane is another couple in the same plane. The moment of the resultant couple must be equal to the algebraic sum of the moments of the component couples.

87. Angular Velocity. — Let θ be the angle through which a body turns in t seconds. Then the mean angular velocity during the t seconds is θ/t. *The angular velocity of a body is*

the time rate of angular displacement. If the angle is expressed in radians and the time in seconds, the quotient θ/t is in radians per second.

Since 2π radians are described in one revolution, if a body makes n revolutions per second, the angular velocity will be

$$\omega = 2\pi n = \frac{2\pi}{T}, \tag{29}$$

where T is the period of rotation.

If Figure 36 is a wheel or a cylinder rotating about an axis through O and making n revolutions per second, its angular velocity is $2\pi n$ radians per second. Any particle m at distance r from the axis describes a circular path of which the circumference is $2\pi r$. Since the particle traces this path n times a second, its *linear velocity* v is $2\pi rn$ units a second. But $2\pi n$ is the angular velocity ω of the wheel in radians per second. Therefore

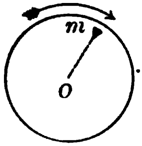


Fig. 36

$$v = r\omega. \tag{30}$$

88. Angular Acceleration. — When an unbalanced torque is applied to rotate a body, such as a flywheel or a cylinder, for example, the effect is to change the angular velocity. There is then an angular acceleration of the body and a linear acceleration of every particle in it. *Angular acceleration is the time rate of change of angular velocity*

If the angular velocity changes from an initial value ω_0 to a value ω in t seconds, the mean angular acceleration α is

$$\alpha = \frac{\omega - \omega_0}{t} \tag{31}$$

in radians per second per second. But linear acceleration a is the rate of change of linear velocity. Therefore

$$a = \frac{v - v_0}{t} = \frac{r\omega - r\omega_0}{t} = r\frac{\omega - \omega_0}{t}$$

and
$$a = r\alpha, \tag{32}$$

or the linear acceleration of any particle in a rotating body is equal to the product of its angular acceleration and its radius of rotation. Linear velocity equals the radius times angular velocity; linear acceleration equals the radius times angular acceleration.

89. Kinetic Energy of Rotation. Moment of Inertia. — A particle of mass m (Fig. 36), situated at a distance r from axis of rotation through O, has, let us suppose, an angular velocity ω. Its linear velocity v is $r\omega$ and its kinetic energy ($\frac{1}{2}mv^2$) is

$$w = \tfrac{1}{2}mr^2\omega^2.$$

The kinetic energy of the entire rotating body is the sum of such expression as this for all the particles composing the body, or

$$W = \tfrac{1}{2}\Sigma mr^2\omega^2.$$

But since the angular velocity ω is the same for every particle of the body, it may be placed outside the sign of summation Σ, and we have

$$W = \tfrac{1}{2}\omega^2\Sigma mr^2. \tag{33}$$

The quantity Σmr^2 is called the *moment of inertia* of the body about the axis through O. It measures the importance of the body's inertia with respect to rotation. The work done on the body to give it an angular velocity ω about any axis is therefore proportional to its moment of inertia about the same axis. With a given angular velocity, the energy of rotation of a body depends not only on its mass, but also on the manner in which that mass is disposed about the axis. The moment of inertia will hereafter be represented by the letter K.

90. Torque and Moment of Inertia. — Referring again to Figure 36, if the particle m has a linear acceleration a directed along the tangent to the circle in which the small mass m revolves, the expression for the tangential force to produce this acceleration is by equation (15) ma. The moment of this force about O is mra. But since a equals $r\alpha$, the mo-

ment of the force on the particle m is $mr^2\alpha$. The sum of the moments of the forces on all the particles composing the body is then $\Sigma mr^2\alpha$. But since the angular acceleration α is the same for all the particles, the expression for the turning moment or torque may be written

$$Fb = \alpha\Sigma mr^2 = \alpha K, \tag{34}$$

where F is the force applied to produce rotation against the resistance due to the body's inertia, and b is its lever arm, or its perpendicular distance from the axis of rotation.

From this expression it is obvious that the moment of inertia may be defined as *the torque required to produce unit angular acceleration;* that is, an angular acceleration of one radian per second per second.

91. Moment of Inertia of a Uniform Rod.—As an illustration of the method of calculating the moment of inertia, let us suppose the uniform rod (Fig 37) to be mounted so as to rotate about an axis through one end O, the axis being perpendicular to the length of the rod. Let the rod be divided into a very large number n equal lengths, and let the mass of each small section be m. Then the whole mass M of the rod is mn. The distances of the several elements of the rod from O are 1, 2, 3, etc., to n in terms of the equal divisions of the rod as the unit of length. The moment of inertia of the rod is the sum of the moments of inertia of the parts, or

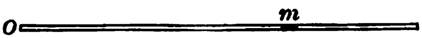


Fig. 37

$$K = m(1^2 + 2^2 + 3^2 + \cdots n^2).$$

Whence $$K = m\frac{n(n+1)(2n+1)}{6}.$$

Since now n is indefinitely large, this expression becomes

$$K = m\frac{n^3}{3} = M\frac{n^2}{3}.$$

But since n is the number of units in the length of the rod, it may be replaced by l, and then finally

$$K = M\frac{l^2}{3}.$$

In a similar way it may readily be shown that the moment of inertia of a uniform rod about an axis perpendicular to its length and passing through its middle point is

$$K_0 = M\frac{l^2}{12}.$$

TABLE OF MOMENTS OF INERTIA

SOLIDS OF MASS M	POSITION OF AXIS OF ROTATION	VALUE OF K
Circle of radius r	Through center perpendicular to plane	$\frac{1}{2} Mr^2$
Cylinder of radius r	Axis of cylinder	$\frac{1}{2} Mr^2$
Hollow cylinder, inner radius r_1, outer r_2	Axis of cylinder	$\frac{1}{2}M(r_1^2 + r_2^2)$
Rectangle, length a, width b	Right angles to plane and through center of figure	$\frac{1}{12} M(a^2 + b^2)$
Sphere of radius r	Through center of sphere	$\frac{2}{5} Mr^2$

92. The Simple Pendulum. — The *simple pendulum* is an ideal pendulum consisting of a heavy particle suspended by a thread without weight. An approximation to a simple pendulum may be made by suspending a small dense sphere from a fixed point by means of a fine thread. If the small sphere is slightly displaced and then released, it will oscillate back and forth about its position of rest. Its excursions become gradually smaller, but if the arc traversed is small, the period of its swing remains unchanged. This characteristic of pendular motion first attracted the attention of Galileo, who noted it in the oscillations of a bronze chandelier suspended by a long rope from the roof of the cathedral in

Pisa. This "lamp of Galileo" may still be seen in the same place.

A *single vibration* is the motion from N (Fig. 38) to either B or E and back again; a *complete* or *double vibration* is the motion from N to B, across to E, and then back again to N. The *period* of a complete oscillation is the interval between two successive passages of the pendulum bob through N in the same direction; that is, it is the period of a complete or double oscillation. The period of a single oscillation is half that of a double oscillation. The *amplitude* is the arc BN or NE.

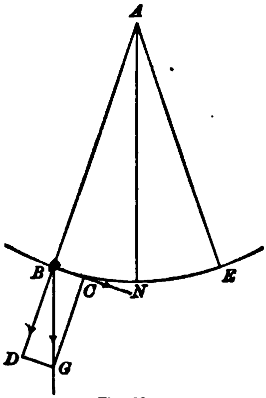


Fig. 38

93. Law of the Pendulum. — When the amplitude does not exceed 2° or 3°, the period of the simple pendulum depends on the length of thread l and the acceleration of gravity g only.

Let AN be a simple pendulum, with a small mass m at the bottom, in its position of rest (Fig. 38). Let it be deflected to the position AB. The acceleration due to gravity g, represented by the vector BG, may be resolved into two components; one, BD, in the direction of the thread, and the other tangential. The former is annulled by the constraint of the thread; the latter is

$$BC = -g \sin \theta,$$

θ denoting the angle BAN, or its equal BGC. Then if θ is small, $\sin \theta$ may be put equal to θ (§ 23). Moreover, $\theta = x/l$, where x is the arc or displacement BN. Therefore

$$BC = -g\theta = -\frac{g}{l}x.$$

But g/l is a constant. Hence the acceleration of B *along the arc* is proportional to the displacement x from the middle

point N. This relation is characteristic of simple harmonic motion. The motion of the pendulum is therefore simple harmonic to the same approximation that $\sin\theta = \theta$.

The acceleration a_x in any direction x for simple harmonic motion is (§ 38)

$$a_x = -a\frac{x}{r} = -\frac{4\pi^2 r}{T^2}\cdot\frac{x}{r} = -\frac{4\pi^2}{T^2}x.$$

Putting this general expression for the acceleration in simple harmonic motion equal to the particular one for the simple pendulum, we have

$$\frac{4\pi^2}{T^2}x = \frac{g}{l}x.$$

Solving for T, $$T = 2\pi\sqrt{\frac{l}{g}}.$$

This is the period of a complete or double swing. For a single oscillation,

$$T = \pi\sqrt{\frac{l}{g}}. \qquad (35)$$

The following are the laws for a simple pendulum:

I. *The period of vibration is independent of the amplitude, if the latter is small.*

II. *The period of vibration is proportional to the square root of the length.*

III. *The period of vibration is inversely proportional to the square root of the acceleration of gravity.*

94. The Physical Pendulum. — It is physically impossible to realize a simple pendulum with its mass all at a single point. Such a pendulum has only an imaginary existence; any real pendulum which does not conform to it is called a *compound* or *physical pendulum.* It is possible, however, to find the length of the corresponding simple pendulum, which will perform its oscillations in the same time as the physical

pendulum. This operation is called finding the length of the equivalent simple pendulum.

Let the mass of the physical pendulum be M, and let its center of inertia be at a distance h from the axis A, about which it oscillates (Fig. 39). The weight of the pendulum is Mg, and this is the force the moment of which produces rotation about the axis A. The lever arm of this force with reference to A as the origin of the moments is $h \sin \theta$, and the turning moment or torque is

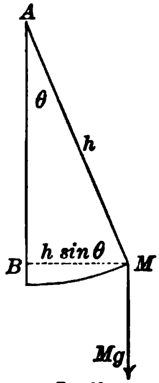


Fig. 39

$$Mgh \sin \theta.$$

But when θ is small, this moment is approximately $Mgh\theta$. By § 90 this moment may also be expressed in terms of angular acceleration and the moment of inertia as αK, or $\frac{4\pi^2}{T^2}\theta K$ (§§ 88, 93.) In both expressions for the moment θ is the small displacement of the center of inertia of the pendulum. Equating the two expressions for the moment of the force,

$$Mgh\theta = \frac{4\pi^2}{T^2}\theta K.$$

Solving for T, $$T = 2\pi\sqrt{\frac{K}{Mgh}}.$$

For a single swing $$T = \pi\sqrt{\frac{K}{Mgh}}. \qquad (36)$$

Comparing this expression with equation (35), it will be seen that the physical pendulum swings as if it were a simple pendulum whose length is K/Mh. This is therefore the length l of the equivalent simple pendulum.

95. The Center of Oscillation. — Let AB (Fig. 40) be a bar suspended so as to swing freely about a horizontal axis through C. C is the center of suspension. Let the center of

mass be at G. Such a physical pendulum has a period of vibration equal to that of a simple pendulum, the length of which is $l = K/Mh$. Lay off the distance l on the line CG produced, so that $CD = l$. The point D is called the *center of oscillation*. The length of the equivalent simple pendulum is therefore the distance between the centers of suspension and oscillation. It follows that if the whole mass of the physical pendulum were concentrated at the center of oscillation, its period as a simple pendulum would be the same as that of the actual physical pendulum.

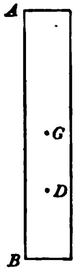


Fig. 40

One of the interesting properties of the center of oscillation is the following: If the pendulum is struck a blow at its center of oscillation, it will be set swinging around its center of suspension, and the blow will not produce any pressure or shock on the axle or knife-edge on which the pendulum is supported. For this reason the center of oscillation is said to coincide with the *center of percussion*.

A baseball bat swung in the hands has a center of percussion, and it should strike the ball at this point to avoid jarring the hands. If the bat is struck higher up or lower down, a distinct "sting" will be felt. If a thin strip of wood about a meter long is held between the thumb and forefinger near one end, and a blow is struck on the flat side with a soft mallet, a point may be found where the blow will not throw the wood strip into rapid vibrations or shivers, but will only set it swinging like a pendulum. This point is the center of percussion or the center of oscillation.

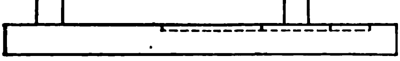


Fig. 41

The position of the center of percussion may be determined experimentally by means of a device shown in Figure 41. A heavy bar is loosely pivoted on a slender axle at S and rests on a blunt edge at P. The position of P is adjustable toward or away from S. If the bar be lifted and allowed to drop on the edge P, the slender pivot at S will break unless P is near the center of percussion of the bar. If P is too near S, the pivoted end will be thrust upwards by the blow; if it is too far away, the thrust will be downwards.

96. The Two Centers Interchangeable. — Huyghens, a celebrated Dutch physicist, discovered that the centers of suspension and oscillation are interchangeable; that is, the period of vibration of the pendulum is the same whether it swings from the one as an axis or the other. This discovery led to the invention by Captain Kater of a pendulum with two parallel axes of suspension, and with weights which can be adjusted until the pendulum has the same period of vibration about the two axes. The distance between them is then the length l of the equivalent simple pendulum.

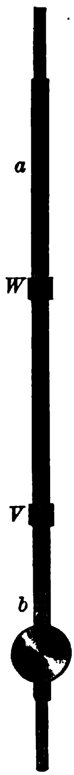


Fig. 42

One form of this pendulum is shown in Figure 42. At a and b are the two knife-edges turned toward each other. W and V are adjustable weights. When the weights have been adjusted so that the period is the same about a as about b, the distance between a and b is the length l of the equivalent simple pendulum.

97. Accelerations of Gravity Compared by Means of the Pendulum. — When the adjustments of a Kater's pendulum have been finally completed, the length l, together with the observed period of vibration T, when inserted in equation (35), will give the value of g at the place of observation.

Moreover, if the period of this same pendulum is observed at different points on the earth's surface, the corresponding accelerations of gravity may be compared, since equation (35) shows that the period is inversely as the square root of g; whence

$$\frac{g'}{g} = \frac{T^2}{T'^2}.$$

98. Length of Seconds Pendulum. — If g is known, the length of a pendulum beating seconds may be found by placing T equal to unity in equation (35) and solving for l. Thus, at New York g equals 980.19 cm./sec.^{2}.

Therefore, $$1 = \pi^2 \frac{l}{980.19}.$$

Whence $$l = 99.3.$$

99. Utility of the Pendulum.—The discovery of Galileo suggested an obvious use of the pendulum as a timekeeper. In the common clock the oscillations of the pendulum regulate the motion of the hands. The train of wheels is kept in motion by a weight or a spring, and the regulation is effected by means of the escapement (Fig. 43). The pendulum rod, passing between the prongs of a fork, communicates its motion to an axis carrying the escapement, which terminates in two pallets. These pallets engage alternately with the teeth of the escapement wheel, one tooth of the wheel escaping from the pallet every double vibration of the pendulum. The escapement wheel is a part of the train of the clock; and as the pendulum controls the escapement it also controls the motion of the hands.

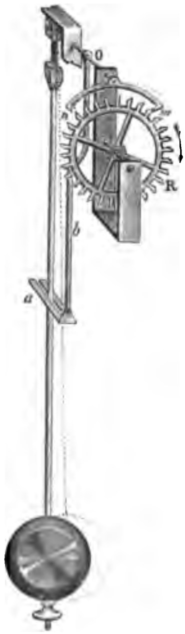


Fig. 43

100. Compensation for Temperature.—Since the vibration period is affected by changes in the length of the pendulum, a common clock with an uncompensated pendulum suffers a change of rate with a change of temperature, losing time in hot weather and gaining in cold. A correction may be made by raising or lowering the bob by means of the running nut on which it rests.

Astronomical clocks for precise measurements have compensated pendulums, which adjust themselves automatically when the temperature changes. The mercurial pendulum, commonly used for this purpose, has in it a mass of mercury, which expands upward while the pendulum rod expands downward, thus effecting a compensation. In one form the mercury in glass tubes forms the pendulum bob; in another the bob is lens-shaped, and the mercury partly fills a steel tube carrying the bob, similar to the one shown in the figure.

VI. EQUILIBRIUM

101. Definition of Equilibrium. — When the forces acting on a material particle have no resultant, that is, when their vector sum is zero, they produce no acceleration and are in equilibrium. It does not follow that the particle is at rest because the forces acting on it are in equilibrium. Equilibrium does not mean that the velocity of the particle is zero, but that its acceleration is zero. Rest means zero velocity; equilibrium, zero acceleration. If the particle is at rest, it will remain at rest; if in motion, it will continue to move without change when the forces acting on it are balanced, or are in equilibrium.

102. Conditions for the Equilibrium of a Particle. — A particle cannot be in equilibrium if it is acted on by a single force only.

Two forces must fulfill the following conditions for equilibrium: they must be equal in magnitude; they must be opposite in direction. These two conditions are sufficient for a material particle because it is incapable of rotation, or at least its rotation has a vanishingly small dynamical significance.

For three forces the condition of equilibrium is that each force must be equal and opposite to the resultant of the other two; for we may replace two of the forces by their resultant, and the problem will then be reduced to the equilibrium of two forces.

This condition may be resolved into two others. It follows, first, that the three forces must lie in the same plane; for the resultant of two of them, as P and Q for example, lies in the plane of P and Q; and then the third force must lie in the same plane, or it could not be opposite to the resultant of P and Q for equilibrium.

Again, it follows that the three forces must be so related to one another that they can be represented in magnitude and direction by the three sides of a triangle, taken in order the

same way round (Fig. 44). If the three forces Q, P, and R can be represented by the sides AB, BC, and CA of the triangle, then R is equal and opposite to R', the resultant of P and Q, and there is equilibrium.

In a similar manner it may be shown that the condition for the equilibrium of any number of forces acting on a particle is that each force must be equal and opposite to the resultant of all the others.

If the forces then all lie in the same plane, they must have such relative magnitudes that they may be represented by

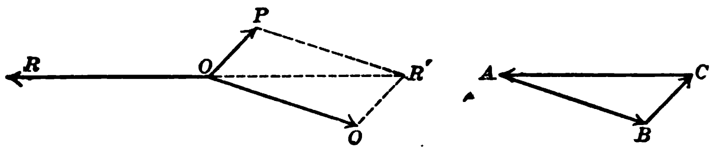


Fig. 44

the sides of a closed polygon taken in order around the figure. If they do not all lie in the same plane, the algebraic sum of their components in each of three rectangular directions must be zero.

103. Conditions for the Equilibrium of a Rigid Body. — The forces applied to a body of sensible dimensions may not all pass through a single point. The necessary condition for equilibrium is therefore that the forces must produce neither translatory nor rotatory acceleration, for such a system of forces may produce either motion of translation, or motion of rotation, or both. A single force applied to a rigid body produces both translatory and rotatory motion, unless its direction passes through the center of inertia of the body.

The first condition of equilibrium is, therefore, that the vector sum of all the forces shall be zero; that is, that the algebraic sum of the components of all the forces taken in three rectangular directions shall be severally equal to zero. There will then be no rectilinear acceleration in any direction.

Again, that the forces may not produce motion of rotation around any axis, it is necessary that the algebraic sum of their moments about any three non-coincident axes shall be zero. The three axes are usually taken at right angles to one another.

If the forces all lie in one plane, it is sufficient for equilibrium that their vector sum be zero, and that the algebraic sum of their moments about any point in the plane be zero.

104. Equilibrium under Gravity. — It is convenient to divide this topic into three divisions:

A. *On a horizontal plane.* The weight of the body is then equal and opposite to the reaction of the plane. Therefore the vertical line through the center of gravity of the body must fall within its base of support. If this vertical falls outside the base, the weight of the body and the reaction of the base form a couple, and the body will overturn.

If the plane on which the body rests is moving up or down with uniform velocity, like the floor of an elevator, there is no acceleration, and the body is in equilibrium, though not at rest. When the elevator starts to ascend, it has acceleration; the pressure of the body on the floor is then greater than its weight. When the elevator starts to descend, it has an acceleration downward, and the pressure of the body on its floor is less than its weight Mg. If the elevator should start with an acceleration equal to g downward, there would be no pressure of the body on the floor of the elevator.

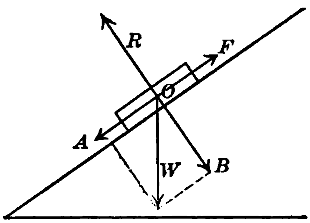


Fig. 45

B. *On an inclined plane.* If the body rests by friction on an inclined plane (Fig. 45), then the component OA of its weight W down the plane is equal and opposite to the sliding friction F. The other component OB of its weight perpendicular to the plane is in equilibrium with the reaction R of the plane. On an

inclined plane, therefore, the weight of a body is in equilibrium with the sliding friction F and the reaction R of the plane. Strictly, the component OA down the plane and the friction F form a couple, and the farther the center of gravity O of the body from the plane, the greater the moment of this couple; but there will be equilibrium so long as the vertical through O does not fall outside the base of the body, and the angle of elevation of the plane is less than the limiting angle of friction.

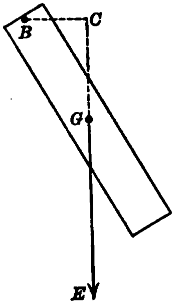


Fig. 46

C. *About a horizontal axis.* If the body is free to turn about a horizontal axis, it can be in equilibrium only when the vertical line through its center of gravity passes through this axis. Let the body whose center of gravity is at G be supported by an axis through B (Fig. 46). Let the vertical line GE represent its weight. When the line GE does not pass through B, the moment of the weight about the axis through B is the product of GE and BC, and the body rotates clockwise. As the body swings, shortening the line BC, the point G approaches the vertical through B; and when the point C coincides with B, the moment becomes zero. The body is then in equilibrium both for rotation and translation.

105. Stable, Unstable, and Neutral Equilibrium. — Consider the case of a body suspended from an axis about which it can turn freely, as in Figure 46. When the body is turned so that its center of gravity G is in a vertical line through B, either above or below, it is in equilibrium because the turning moment is zero. In the second position, with G vertically below B, if we suppose the body slightly displaced about the axis, as in the figure, the moment of the weight of the body about the axis of suspension tends to *decrease* the displacement and to bring the body back again to the position of equilibrium. The equilibrium is then said to be *stable*.

If on the contrary the center of gravity G be vertically above the axis B, and the body be slightly displaced, the turning moment will tend to *increase* the displacement. Hence the equilibrium in which the center of gravity of the body is vertically above the axis of suspension is *unstable.*

If the axis of suspension passes through the center of gravity of the body, a displacement of the body does not bring into operation a couple tending either to increase or decrease it, and the equilibrium is *neutral.*

The three kinds of equilibrium may be illustrated also by a body resting on a horizontal plane. Let a cone (Fig. 47) rest on a horizontal plane in the position A. It is in stable equilibrium. In the position B it is in unstable equilibrium, which is a physically impossible equilibrium. In the position C, its equilibrium is neutral for a rolling displacement. A ball on a horizontal plane is in neutral equilibrium for any point on its surface in contact with the plane.

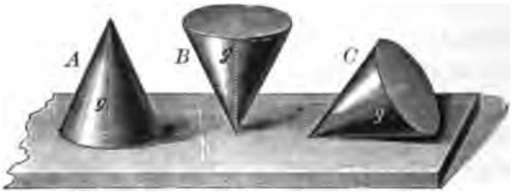


Fig 47

The three kinds of equilibrium for a body are therefore these : *stable*, for any displacement which causes its center of gravity to rise ; *unstable*, for any displacement which causes its center of gravity to fall ; *neutral*, for any displacement which does not change the height of its center of gravity.

In general, if a body in neutral or unstable equilibrium is slightly displaced, it has no tendency to return, and consequently will not rock about its position of neutral or unstable equilibrium. An oscillation, then, is always a motion around or through a position of stable equilibrium.

106. Stability.—The most useful measure of the stability of a body is the work necessary to overturn it; that is, it is

the product of its weight and the difference between the distances AC and AD in Figure 48. In the diagrams C is the

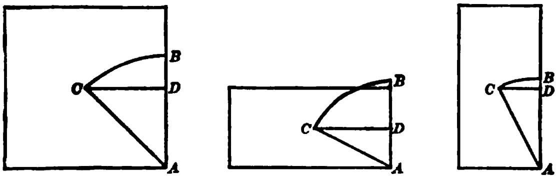


Fig. 48

center of gravity, A the point about which the body is turned, and BD the height through which the body is lifted to bring it to a position of unstable equilibrium.

A brick has less stability when standing on end on a level surface than when lying on edge; and it has less stability on edge than when lying on its broad side.

Suppose a short cylinder of wood loaded on one side, so that its center of gravity falls at G (Fig. 49). This cylinder may be placed on a plane,

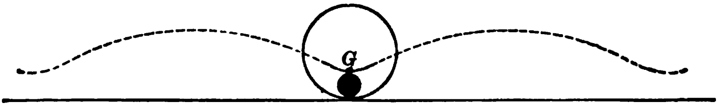


Fig. 49

slightly inclined, in such a position that it will roll up the plane into a position of stable equilibrium.

If the cylinder be rolled along a horizontal plane, its center of gravity will describe a curve with crests and hollows, similar to the one shown in the figure. Every hollow corresponds to a position of stable equilibrium, and every crest to one of unstable equilibrium. If the path were a straight line, the equilibrium would be neutral.

VII. MACHINES

107. Definition of Machines. — A *machine* is a device designed to change the direction or the magnitude of the forces required to accomplish some useful purpose, or to transform and transfer energy. For example, by the use of a simple pulley to change the direction of the force, a body may be *lifted* while the force is applied in a *downward* direction.

Again, by means of a bicycle a man propels himself forward, or even up an incline, by a downward and backward thrust of his feet on the pedals.

In the approaches to the Saint Gothard tunnel through the Alps, an ascent insurmountable directly is made possible by extending the elevation to be made over a long inclined plane, like the thread of a screw, cut partly in the face of the mountain and partly in circular tunnels through the rock. In this manner the force required to lift the train is reduced to the capacity of the locomotive. The effect of this engineering device is to reduce the rate of doing the work of lifting the train.

A steam engine is a machine designed to transform the heat energy applied to it into useful mechanical work; a dynamo-electric machine transforms mechanical energy into the energy of an electric current. A water wheel transforms the potential and kinetic energy of falling water into mechanical energy represented by the shaft of the wheel turning with a definite torque.

A complex machine, consisting of a train of mechanism, may comprise a series of *simple* or *elementary machines*, or *mechanical powers*. These are: the *lever*, *pulley*, *wheel and axle*, *inclined plane*, and *screw*. All complex machines are mechanically only combinations of two or more simple machines.

108. Mechanical Advantage.—All machines designed to transform or transfer energy have the common characteristic that the force applied at one part to produce motion enables work to be done at another part against resistance. The force applied is called the *effort*, and the force worked against, the *resistance*. The problem in simple machines is reduced to finding the ratio of the resistance to the effort. This ratio is taken as the measure of the *mechanical advantage* of the machine.

It is evident that the advantage derived from the use of machines is not all mechanical; for in many cases the advantages gained in other respects may more than compensate for a loss of technical mechanical advantage.

In elementary discussions it is customary to neglect friction and to assume that the parts of a machine are rigid and without weight.

109. General Law of Machines.—Every machine must conform to the principle of the conservation of energy; that is, *the work done by the effort must equal the work done in overcoming the resistance*, except that some of the energy may be dissipated as heat or may not appear in mechanical form. A machine can never produce an increase in the amount of energy applied.

Denote the effort by F and the resistance by R, and let d and D denote the distances through which they work respectively. Then from the law of conservation of energy,

$$Fd = RD, \tag{37}$$

or the effort multiplied by the displacement in its direction is equal to the resistance multiplied by the displacement directly against it.

110. Efficiency.—If all wasteful resistance could be eliminated from a machine, it would waste no energy and its efficiency would be unity. But in practice there is always present some wasteful resistance due to friction, rigidity of cords, etc. The work done is, therefore, always partly useful and partly wasteful. The efficiency of a machine is the ratio of the useful work done by it to the total work done on it; it is the output divided by the input of energy. Efficiency is always a proper fraction and it is expressed as a percentage. An efficiency of 90 per cent means that the energy recovered is 90 per cent of the energy put into the machine. A machine that will do either useful or useless work continuously, without receiving a continuous or intermittent supply of energy from without, is clearly an impossibility.

111. The Lever.—A *lever* is a rigid bar turning about a fixed axis called the *fulcrum*. The perpendicular distances between the fulcrum and the lines of action of the effort and

the resistance are called the *arms* of the lever. A straight lever has its two arms in the same straight line.

Several cases arise according to the relative positions of the forces with respect to the fulcrum. If the fulcrum is between the effort and the resistance, the lever is of the *first* kind; if the resistance is between the effort and the fulcrum, the lever is of the *second* kind; and, finally, if the effort is between the resistance and the fulcrum, the lever is of the *third* kind. The three kinds of lever are represented by the three diagrams of Figure 50 in order.

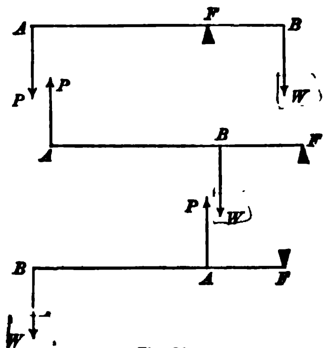


Fig. 50

A crowbar is a lever of the first kind. A pair of scissors consists of two levers of the first kind joined together. A nutcracker is a double lever of the second kind. A pair of spring shears, used for shearing sheep or clipping grass, is a double lever of the third kind. So is an ordinary pair of tongs. The forearm is also an example of a lever of the third kind; the fulcrum is at the elbow, the resistance is at the hand, and the effort is applied by the biceps muscle between the two.

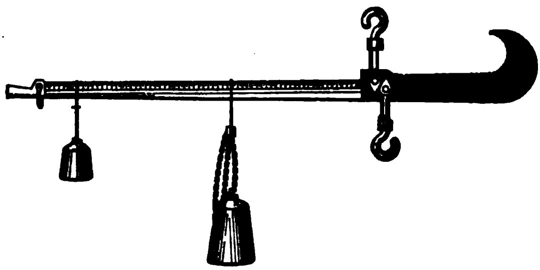


Fig. 51

The *steelyard* (Fig. 51) is a lever of the first kind with unequal arms. It is a form of balance in which the body to be weighed is suspended from the shorter arm of the lever, and a counterpoise is caused to slide along the longer arm to produce equilibrium. The center of gravity of

the steelyard is at the fulcrum. The body weighed is heavier than the counterpoises; the divisions of the beam are equidistant and indicate the weight.

112. Mechanical Advantage of the Lever.—In the three diagrams of Figure 50, the arms of the lever are AF and BF. Let P be the effort and W the resistance (a weight or other force). The most direct and simple way to obtain the relation between the effort and resistance is to take moments around the fulcrum as the origin. If the lever is to be in equilibrium, these moments must be equal and opposite. Hence

$$P \times AF = W \times BF, \text{ or } \frac{W}{P} = \frac{AF}{BF}.$$

Therefore, *the mechanical advantage of the lever equals the inverse ratio of the arms.*

If it is desired to take into account the weight of the lever, the moment of this weight, considered as acting at the center of gravity of the bar, must be added to either the moment of the effort or of the resistance, according as the weight acts to produce rotation in the direction of the one or the other.

In this discussion it has been tacitly assumed that the forces act at right angles to the length of the bar. If they do not, then the lever arms are the distances measured perpendicularly from the fulcrum to the directions of the two forces.

113. The Wheel and Axle. — This simple machine may be considered as a continuous lever. It consists of two cylinders of different diameters turning together on the same shaft. The center of the shaft is at C (Fig. 52). A rope is wound around each cylinder, right-handedly around one and in the opposite direction around the other. When the cylinders are turned, one rope unwinds and the other winds up.

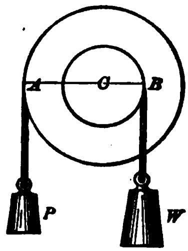


Fig. 52

If the radii of the large and the small cylinders are R and r respectively, during one revolution the distance worked through by the effort P is $2\pi R$, and the distance moved against the resistance W is $2\pi r$. Hence

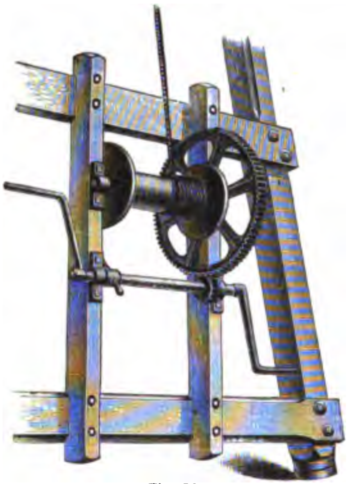


Fig. 53

$$2\pi R \times P = 2\pi r \times W.$$

Whence $\frac{W}{P} = \frac{R}{r}$.

The weight P may be replaced by any effort P applied to the circumference of the wheel, and the weight W by any resistance W at the circumference of the axle. *The mechanical advantage is the ratio of the radius of the wheel to that of the axle.*

114. Application. — The *derrick* is a form of wheel and axle much used for lifting heavy weights in construction work. The essential parts are shown in Figure 53. The derrick may be considered as two sets of wheel and axle in series. The axle of the first set works on the wheel of the second by means of the spur gears. The cranks of the first set answer the same purpose as a wheel. The mechanical advantage is the product of the radii of the wheels divided by the product of the radii of the axles.

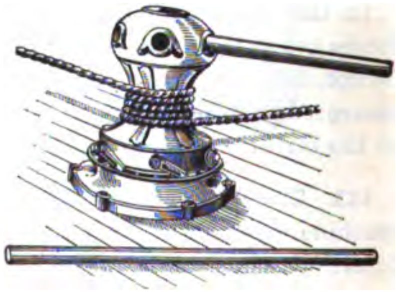


Fig. 54

In the *capstan* (Fig. 54) handspikes inserted in holes at the top are used instead of a wheel; the rope by which the work is done is wrapped around the body of the capstan as an axle.

115. The Pulley. — The pulley is a wheel, called a *sheave*, free to turn about an axle in a frame called a *block*. The

effort and the resistance are connected by means of a rope, which lies in a groove cut in the circumference of the wheel. The object in using the wheel instead of a fixed cylinder to change the direction of the rope is to reduce friction.

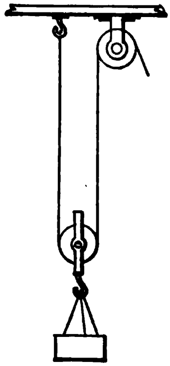


Fig. 55

In the simple movable pulley of Figure 55 the effort, or tension in the rope, is half the weight or other resistance. But the lower sheave is supported by two branches of the rope. If the weight is lifted, it rises half as fast as the free part of the rope travels. The mechanical advantage of such a pulley with one movable block is obviously equal to two.

The most useful pulley consists of two blocks, each with several sheaves turning on the same axle. One of the blocks is fixed, while to the other is attached the weight or other resistance (Fig. 56).

Fig. 56

The principle involved in determining the mechanical advantage of a system of pulleys is the transmission of the same tension to all parts of the rope. In reality the rope is stiff and there is friction at the axles. The effect of this rigidity and friction is a diminution in the tension of the rope as it passes a pulley. If this diminution is a fixed ratio, allowance can be made for it.

When there is one continuous rope passing around the two pulleys, as in Figure 56, it is obvious that the weight is sustained by the several parts of the rope, the tension in each part being the effort P applied at the free end, neglecting rigidity and friction. If there are n lengths of the rope between the fixed and movable blocks, the sum of the tensions supporting the

weight or resistance is nP; that is,

$$W = nP, \text{ and } \frac{W}{P} = n.$$

When a single rope is used, *the mechanical advantage of a system of pulleys is therefore equal to the number of times the rope passes between the two blocks.*

116. The Inclined Plane.— Suppose a body resting on an inclined plane without friction. The weight of the body acts vertically downward, while the reaction of the plane is perpendicular to its surface; so that a third force must be applied to maintain the body in equilibrium on the incline. Two principal cases occur: first, when the force is applied parallel to the *face* of the plane; second, when it is applied parallel to the *base* of the plane.

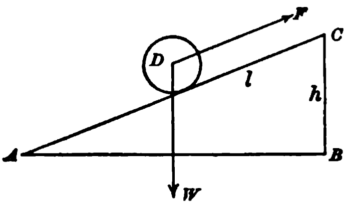


Fig. 57

A. The most convenient method of obtaining the relation between the force F (Fig. 57) and the weight W is to apply the principle of work. Suppose D to move under the influence of the force F from A to C without acceleration up the plane. Then the work done by the force F is $F \times AC$. The work done on the body D against gravity in lifting it through the vertical height BC is $W \times BC$, and

$$F \times AC = W \times BC,$$

or

$$\frac{W}{F} = \frac{AC}{BC} = \frac{l}{h}.$$

The mechanical advantage when the force is applied parallel to the face of the plane is the ratio of the length of the plane to its height.

B. When the force is applied parallel to the base of the plane, the component of the displacement in the direction of

the force, when the body D moves from A to C (Fig. 58), is the base of the plane AB. Therefore the work done by F is $F \times AB$. The work done on the weight W against gravity is the same as in the first case.

Hence

$$F \times AB = W \times BC,$$

or

$$\frac{W}{F} = \frac{AB}{BC} = \frac{b}{h}$$

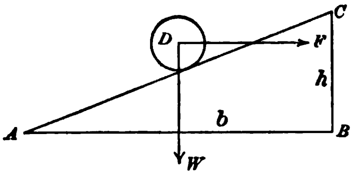


Fig. 58

The mechanical advantage when the force is applied parallel to the base of the plane is the ratio of the length of the base of the plane to its height.

117. The Screw. — The screw is a cylinder on the surface of which is a uniform spiral called the *thread*. The faces of the thread are inclined planes. If a long triangular strip of paper be wound on a cylinder such as a pencil, with the base of the triangle perpendicular to the axis of the cylinder, the hypotenuse of the triangle will form the thread of a screw (Fig. 59). The distance s between successive turns of the thread is called the *pitch* of the screw.

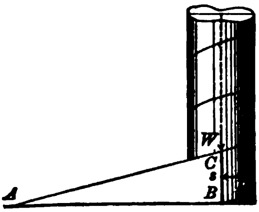


Fig. 59

The screw works in a block called a *nut*, on the inner surface of which is a spiral groove. This groove is the exact counterpart of the thread (Fig 60). When either the screw or the nut makes a complete turn, the relative motion of the two parallel to the axis of the screw is the distance s.

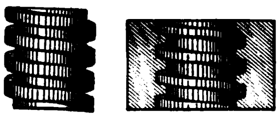


Fig. 60

The screw is usually combined either with a lever or its equivalent, a wheel. The mechanical advantage may then be found most readily by applying the principle of work expressed in the general law of machines (§ 109). If s is the pitch of the screw, and l is the

lever arm or radius of the wheel, then when the effort P makes a complete turn, the equation of a work is

$$P \times 2\pi l = Ws,$$

or

$$\frac{W}{P} = \frac{2\pi l}{s}.$$

Hence, *the mechanical advantage of the screw is the ratio of the distance traversed by the effort in one turn to the pitch of the screw.*

In deducing the above relation the friction between the screw and the nut has been disregarded. This friction is always far from negligible, and in practice the mechanical advantage is considerably less than $\frac{2\pi l}{s}$.

118. The Screw Gauge. — The screw in the form of a *screw gauge* is used for measuring small dimensions. The object to be measured is placed between the end of the screw C and the block B (Fig. 61). The nut is held on the other end of the strong curved arm. The head of the screw, or the cap D, is divided into some number of equal parts, say 50. The whole number of turns made by the screw is read on the scale A, which is uncovered by the movement of the cap attached to the screw. The fractions of a turn are read on the scale on the edge of the cap. If, for example, the pitch of the screw is $\frac{1}{2}$ mm., then for each turn the end of the screw moves $\frac{1}{2}$ mm.; and if the scale on the cap reads 12 divisions further, the screw has moved $\frac{12}{50} \times \frac{1}{2} = 0.12$ mm. in addition.

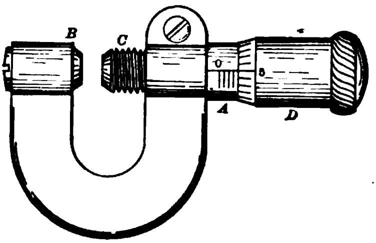


Fig. 61

119. Sensibility of the Balance. — The balance is an instrument for the comparison of equal masses. It is essentially

a lever of the first kind with arms of equal length. It consists of a light, trussed beam, so as to have the requisite stiffness with the least weight. It is supported at its middle point by means of a knife-edge resting on agate planes. The scale pans are of equal weight and are suspended on knife-edges from the ends of the beam.

A good balance must fulfill the following conditions: (1) It must be true; that is, the beam must be horizontal with equal weights in the two scale pans. (2) It must be sensitive; that is, a small difference between the two masses in the scale pans must produce an observable deviation of the beam from a horizontal position. (3) It must be stable; that is, the beam must return to its horizontal position of equilibrium very precisely after displacement. (4) A fourth desideratum is that its period of oscillation about its position of stable equilibrium shall be as small as possible.

Let the three points A, B, and C (Fig. 62) be in the same straight line. A and B are the knife-edges for the support of the pans, and C is the knife-edge on which the beam rests. Let G be the center of gravity of the beam the weight of which is w. Let a weight P be placed in one scale pan and $P+p$ in the other. If then the two arms of the balance are of equal length, P and P in the two pans have equal moments about C and counterbalance each other. They may then be omitted from the equation of equilibrium. Then the moment of the small excess weight p about C is equal and opposite to the moment of the weight of the beam w about the same point. Therefore

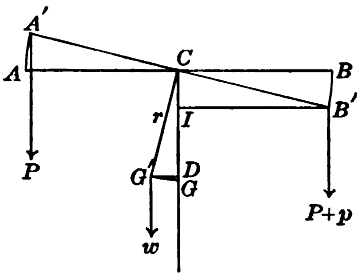


Fig. 62

$$p \times B'I = w \times G'D.$$

G' is the position of the center of gravity of the beam when it is displaced by the excess weight p to the new position $A'B'$. Let CG be denoted by r, and let l be the length of either arm of the balance. Then, if θ is the angular displacement BCB',

$$B'I = l \cos \theta ; \quad G'D = r \sin \theta.$$

Substituting in the preceding equation of equilibrium we have

$$p \times l \cos \theta = w \times r \sin \theta.$$

Whence $$\frac{\sin \theta}{\cos \theta} = \tan \theta = \frac{pl}{wr}. \qquad (38)$$

The sensibility of the balance is measured by the angular displacement θ of the beam with a given small difference of load p. If now θ is small, it is equal to $\tan \theta$ (§ 23), and the sensibility for a given excess weight p is directly proportional to the length l of the balance arm, inversely proportional to the weight w of the beam, and inversely proportional to the vertical distance r of its center of gravity below the knife-edge.

The conditions of sensibility do not agree in every particular with the requirements of a good balance. Especially is it true that a long beam l for high sensibility is directly antagonistic to a short period of oscillation; for the longer the beam, the larger its moment of inertia and the longer its period. Then, too, a long beam means a larger weight w to secure sufficient rigidity; and the sensibility is inversely as the weight of the beam.

In the best modern balances the beam has been shortened, in order to secure lightness of the movable system and a short period of oscillation. At the same time high sensibility has been secured by better workmanship on the knife-edges and bearing planes, and by the use of aluminum alloys to further reduce the weight of the beam.

The bending of the beam under a load raises the point C with respect to the line through A and B. On this account an increase of the load generally produces a decrease in the sensibility.

VIII. ELASTICITY

120. Strain and Stress. — When a body is forced to change its size or shape, it is said to be *strained*, and the deformation it undergoes is called a *strain*. In general, a body resists a strain, and the internal restoring force, tending to cause the body to revert to its unstrained state, is called a *stress*. When the strained body is in equilibrium, the external deforming force is equal and opposite to the stress evoked, and the deforming force may then be called a stress (§ 51). The measure of the stress is the force per unit area. In the *c.g.s.* system a stress is measured in dynes per square centimeter.

121. Kinds of Stress. — When the effect of a stress on a section of a body to which the stress is applied is to increase the dimensions of the body at right angles to the section, the stress is a *tension;* when the effect is to diminish this dimension, the stress is a *pressure*. A stress which alters the form but not the size of a body is called a *shearing stress*. The deformation which a body undergoes under a shearing stress may be aptly illustrated by the aid of a pack of cards, lying on a table and forming a rectangular parallelopiped. Imagine a horizontal force so applied as to cause each card to slip forward over the next one below it by the same amount (Fig. 63). Each card will then move forward a distance proportional to its height above the table, and the pack has undergone a shear.

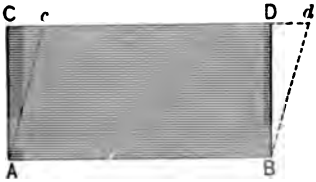


Fig. 63

A stress is called *hydrostatic pressure* when the pressure at a point is the same in all directions. For measuring hydrostatic and gas pressure, the column of mercury which it will support is often used. The unit of measure is then the pressure at a depth of one centimeter in mercury. A cubic centimeter of mercury weighs 13.596 gm. Hence the value of this unit of pressure is $13.596 \times 980 = 13,324$ dynes.

The pressure of illuminating gas is measured in inches or centimeters of water. A column of water one centimeter high produces a pressure per square centimeter of one gram, or 980 dynes.

122. Elasticity. — Solid bodies react against a change either in shape or volume; fluid bodies react against a change in volume only. The property of a body by virtue of which it exerts such a reaction and tends to recover from a strain is called *elasticity*.

A body which exhibits reaction under a change of volume has *elasticity of volume;* a body which reacts against change of shape, but without change of volume, has *elasticity of form*, or possesses rigidity.

If a body suffers a strain, which does not change so long as the stress remains the same, and which completely disappears when the stress is removed, the body is said to be *perfectly elastic*. Gases and liquids have perfect elasticity of volume; that is, they recover perfectly their initial volume when the initial pressure is restored.

There is a limit to the elasticity of solids, called the *elastic limit*, beyond which they yield and are then incapable of regaining their original form and volume. The elastic limit for steel is high, and it breaks before there is much permanent distortion. On the other hand, lead scarcely recovers completely from any distortion, however slight.

123. Elastic Fatigue. — Even within the limits of elasticity, solids show distinct differences in their behavior. Some recover at once after the removal of the force of distortion. A *fine* thread of spun quartz recovers immediately from a twist after the torsional force is removed, while a steel wire may not recover completely for several hours, if it has been kept under torsion for some time. This delay in recovery from distortion is said to be due to *elastic fatigue*. It is very noticeable in the case of glass, and metals are never free from it. If a long glass fiber is kept twisted for some time, it will

largely untwist as soon as it is released from torsion, and it will then creep slowly back to its original condition. The larger the initial twist and the longer it lasts, the greater is the temporary set of the fiber.

124. Hooke's Law. — When the strain in an elastic body does not exceed the elastic limit, *the reaction,* due to the strain and tending to restore the body to its unstrained condition, *is proportional to the distortion.* This relation is known as *Hooke's law.* It has been verified for most materials in common use.

It follows from Hooke's law that if we know the stress corresponding to unit strain of any type, we can find the stress corresponding to a strain of any magnitude within the elastic limit and of the same type. According to this law, stress and strain are connected by the following relation:

$$\textit{Stress} = e \times \textit{strain}, \text{ or } e = \frac{\textit{stress}}{\textit{strain}}.$$

The quantity e is a proportionality factor called the *modulus of elasticity.* It is *the quotient of stress by strain.*

125. Modulus of Volume Elasticity. — If the strain is due to a change in the size of the body only, the measure of the strain is the diminution suffered by unit volume of the strained body; and e becomes the *modulus* or *coefficient of volume elasticity.*

Let the body be subjected to a uniform normal pressure p over its entire surface. Let V be the original volume, and v the diminution in volume. Then v/V is the compression, and this is the measure of the strain. The modulus of volume elasticity becomes

$$k = p \div \frac{v}{V} = \frac{p}{v} V. \qquad (39)$$

This is the only modulus or coefficient of elasticity applicable to liquids and gases.

126. Young's Modulus. — When a wire, for example, is stretched by a weight, the stress is the force per unit of cross section of the wire, the strain is the increase in unit length of the wire, and e is then called *Young's modulus.*

Let L be the unstrained length of the wire, $L + l$ its length when stretched within the limit of elasticity. The strain is l/L. If F is the stretching force and A the sectional area of the wire, F/A is the stress per unit area. Therefore Young's modulus is

$$y = \frac{F}{A} \div \frac{l}{L} = \frac{FL}{Al}. \tag{40}$$

127. Energy stored in a Strained Body.—The work required to stretch an elastic wire or slender rod within the limit of elasticity is stored in the strained body as potential energy. The expression for this energy may be derived from the case of a stretched wire.

Let it be assumed that the load is added so gradually that it does not acquire appreciable velocity, so that none of the work done becomes kinetic energy. Since Hooke's law applies, the force for each increment of the elongation is proportional to the whole elongation corresponding. Then if L is the initial length of the wire, l the total elongation produced by the force F, the mean working force is $\frac{1}{2}F$, and the work done is

$$W = \tfrac{1}{2}\, Fl.$$

But from the definition of stress and strain,

$$F = \mathit{stress} \times A, \quad \text{and} \quad l = \mathit{strain} \times L.$$

The work done is therefore

$$W = \tfrac{1}{2}\, AL \times \mathit{stress} \times \mathit{strain}.$$

But AL is the volume of the wire, and therefore the energy stored in unit volume of the stretched wire is

$$W = \tfrac{1}{2}\, \mathit{stress} \times \mathit{strain}.$$

Problems

1. What is the acceleration when a force of 40 dynes acts on a mass of 5 gm.? How far will the mass move in 4 seconds?

2. A force of 60 dynes acts on a body for one minute and gives to it a velocity of 1200 cm. a second. What is the mass of the body?

3. A mass of 500 gm. is whirled around at the end of a string 40 cm. long twice a second. What is the tension in the string in dynes, disregarding gravity?

4. An inelastic mass of 500 kgm., moving with a velocity of 30 m. a second, meets a similar mass of 300 kgm., moving with a velocity of 20 m. a second in the opposite direction. Find the velocity of the entire mass after impact.

5. Compare the kinetic energy of a ball having a mass of 15 gm. and a velocity of 500 m. a second with that of a gun from which it was fired, if the mass of the gun is 10 kgm.

6. A force of 1500 dynes acts continuously on a mass of 10 gm. for 30 sec. Find the velocity acquired and the space traversed in the 30 sec.

7. A shot weighing 20 lb. is fired from a gun weighing 5 tons with an initial velocity of 1500 ft. a second. What is the initial velocity of recoil of the gun?

8. A rapid-firing gun fires 300 bullets of 30 gm. each per minute with a velocity of 300 m. a second. Find the mean force of reaction on the gun.

9. What angular velocity in a vertical plane must be given to an open vessel containing water so that no water may be spilled?

10. A man weighing 150 lb. climbs to the top of the Eiffel Tower, height 984 ft. How many foot-pounds of work does he do?

11. How many ergs of work are done in raising a mass of 1 kgm. vertically through a height of 5 m.?

12. An unbalanced force of 10 kgm. moves a mass of 100 kgm. through the distance of 100 m. How much work is done?

13. By means of a force of 1000 dynes, a mass is moved a distance of 200 m. How many joules of work are done?

14. How much work in foot-pounds is done against gravity in hauling a load of 2000 lb. to the top of a hill 200 ft. high? If the hill is 2000 ft. long, what force *against gravity* is necessary to pull the load up the hill?

15. At what rate is an engine working which raises 1000 tons of coal in 10 hr. from a mine 300 ft. deep?

16. A steam pump fills a tank with water in 4 hr. The capacity of the tank is 5000 gal. and the elevation is 40 ft. If a gallon of water weighs 8 lb., what is the horse power of the pump?

17. An electric motor raises an elevator cage whose unbalanced weight is 2000 kgm. through a height of 40 m. in 40 sec. What is the power of the motor in kilowatts?

18. How many horse powers are transmitted by a rope passing over a pulley 16.5 ft. in circumference and making one revolution a second, the tension in the rope being 200 lb.?

19. A dynamometer pulley is 32 cm. in diameter and the cord 1 cm. The weight at one end to produce tension is 1760 gm., and when the pulley makes 720 revolutions per minute, the tension of the cord shown by the spring S (Fig. 31) is 14 kgm. What is the power absorbed by the dynamometer in kilowatts and in horse powers?

20. Two men carry a weight of 150 kgm. slung on a pole 280 cm. long. If the weight be placed at the distance of 100 cm. from one end, what portion of the weight does each man carry?

21. Two cylinders of the same uniform material, each 30 cm. long, and of diameters 12 cm. and 8 cm. respectively, are joined end to end so that their axes are in the same straight line. Where is their common center of gravity?

22. A circle 20 cm. in diameter has cut out of it a smaller circle tangent to it and 12 cm. in diameter. Where is the center of gravity of the remainder?

23. A cylinder whose mass is 2000 gm. and radius 10 cm. rotates on its axis 300 times a minute. Find its kinetic energy of rotation.

24. A pendulum beating seconds at one place is carried to another station where it gains 10 sec. a day. Compare the accelerations of gravity at the two places?

25. A uniform slender rod 1 m. long is pivoted at one end so as to swing as a pendulum. Calculate its period of vibration at a place where g is 980 cm. per second per second.

CHAPTER IV

MECHANICS OF FLUIDS

I. MOLECULAR PHENOMENA IN LIQUIDS

128. Characteristics of a Fluid. — A fluid has no shape of its own, but takes the shape of the containing vessel. It cannot resist a stress unless it is supported on all sides. A perfect fluid would offer no resistance to a shearing stress. The molecules of a fluid at rest are displaced by the application of the slightest force; that is, a fluid yields to the continued application of a force tending to change its shape. Nevertheless fluids exhibit wide differences in *mobility*, or readiness in yielding to a shearing stress. Alcohol, gasoline, and sulphuric ether are examples of very mobile liquids; glycerine is very much less mobile. In fact, liquids shade off gradually into solids, and there are intermediate bodies which exhibit to some degree the properties of liquids as well as of solids. A stick of sealing wax supported at its ends yields continuously, though very slowly, to its own weight. A cake of shoemaker's wax on water, with bullets on it and corks under it, yields to both and is traversed by them in opposite directions. The wax will flow very slowly down a tortuous channel. It is therefore mobile, and its mobility increases with its temperature. At the same time both sealing wax and shoemaker's wax when cold break readily under the blow of a hammer like a solid.

129. Viscosity. — The resistance of a fluid to flowing under stress is called *viscosity*. It is due to molecular fluid friction. The slowness of the descent of a fine precipitate in water is

due to the viscosity of the liquid ; and the slowness of the fall of fine raindrops, or a cloud, is due to the viscosity of the air. Viscosity varies through wide limits. It is less in gases than in liquids, and in general decreases as the temperature rises. Hot water is less viscous than cold water; hence the relative ease with which a hot solution filters.

A body set oscillating in air has its vibrations damped by the molecular friction of the air, that is, by its viscosity ; if suspended in a liquid, the damping is much more pronounced. If the body is wetted by the liquid, the damping is due entirely to viscosity and is independent of the nature of the suspended body ; for there is no loss of velocity by friction between the solid and the liquid.

A perfect fluid would be entirely without rigidity and viscosity.

130. Liquids and Gases. — Fluids include both liquids and gases. The two may be distinguished by two characteristic properties :

First. Liquids have a free surface, while gases cannot permanently retain a free bounding surface, independent of the containing vessel. A gas introduced into an empty vessel completely fills it, whatever its volume.

Second. Liquids are but slightly compressible, while gases are highly compressible, and tend to expand to an indefinitely large volume. A liquid offers great resistance to a stress tending to reduce its volume, while a gas offers relatively small resistance. Both have perfect elasticity of volume, but their coefficients of elasticity differ greatly. Water, for example, is reduced in volume only 0.00005 by a pressure of one atmosphere (§ 171), while air is reduced to one half by the same increase of pressure above that of the atmosphere.

131. Cohesion in Liquids. — In liquids the molecules are within the sphere of one another's attraction. This attraction accounts for the viscosity of even the most mobile liquids. A liquid is hindered in its flow by molecular friction. Molec-

ular attraction accounts for the fact that a small stream of liquid has a certain tenacity and does not break readily.

If a clean glass rod be dipped in water and then withdrawn, a drop will adhere to the end of the rod until enough water has run down the rod to increase the weight of the drop to a point where it falls as a little sphere of water. Its spherical form is due to the attraction between its molecules, which gives to it uniform molecular pressure and a minimum surface.

Cohesion in a liquid is due to the attraction existing among its molecules. It is to be noted that this attraction acts only at insensible distances. It diminishes rapidly as the distance increases and vanishes at a range something like the twenty thousandth of a millimeter.

132. Phenomena at the Surface of a Liquid.—Bubbles of gas formed in the interior of a cold liquid rise to the surface and often show some difficulty in breaking through. A sewing needle carefully placed on water floats. The surface of the water around the needle is depressed and the needle rests in a little hollow. Let two bits of wood float on water a few millimeters apart. When a drop of alcohol is applied to the surface between them, they suddenly fly apart. A thin film of water may be spread over a chemically clean glass plate; but if a drop of colored alcohol falls on this film, the film will break, the water retiring and leaving a dry area around the alcohol.

The sewing needle indents the surface of the water as if this surface were a tense membrane or skin, and tough enough to support the needle. This membrane is weaker in alcohol than in water; hence the moving apart of the bits of wood and the withdrawal of the water from the spot weakened by alcohol.

133. Surface Tension.—It is apparent that the surface of a liquid is physically different from the interior. The molecules composing the surface film are not under the same conditions of molecular equilibrium as those in the interior of the liquid. Let ϵ be the range of molecular attraction. Then at any point in the interior of the liquid, at a distance

from the surface greater than ϵ, each molecule is attracted equally in all directions. But at or very near the surface the attraction downward is not balanced by an equal attraction upward, and in consequence the molecules along the surface are crowded together so as to form what may be considered an elastic film.

Let mn (Fig. 64) be the surface of the liquid, and let $m'n'$ be a parallel plane at a distance ϵ below the surface. For any point P above the plane $m'n'$ the attraction downward is greater than the attraction upward. About P as a center describe a sphere with a radius ϵ. Then the normal pressure on a plane through P perpendicular to mn is greater than on a plane parallel to mn. The upward attraction on P varies from a maximum at $m'n'$ to zero at the surface. As P rises, the upper half of the sphere described about P contains a diminishing number of molecules, but horizontally the attraction remains unchanged. From this inequality there arises a stress, causing the surface to contract.

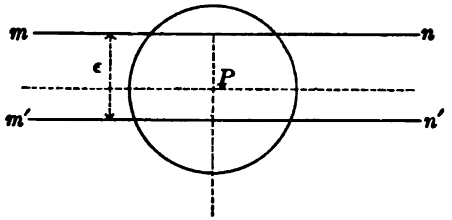


Fig. 64

If we imagine the surface to be enlarged by forcing molecules out along the plane through P normal to the surface, then work must be done on them to transfer them from the interior and to spread them out against the force pressing the molecules together along the surface. It follows that an increase in the surface means an increase in potential energy. A surface film therefore possesses surface energy; and as potential energy always tends toward a minimum, the surface contracts to as small dimensions as possible.

134. Shape of a Small Liquid Mass.—A small mass of free liquid always tends to assume a spherical form, because the

volume inclosed by a sphere is a maximum relative to the area of its surface. The smaller the mass of the liquid, the larger its surface in proportion to its weight, and hence the smaller the influence of gravity in distorting it from the spherical form. Small globules of mercury on clean glass approach the nearer to a spherical form the smaller the globules. Drops of rain and dew are nearly spherical because of surface tension.

When molten lead flows from a small orifice, the surface tension around the small stream throttles it and cuts it into segments; again, surface tension molds these small detached masses into spherical form as soon as their oscillations have died out by the internal friction due to viscosity. If they rotate as they descend, they remain quite spherical and strike the water at the bottom of the shot tower as solid shot.

An ingenious method of separating the perfect shot from the imperfect ones consists in causing all to roll together down a smooth inclined plane. Near the bottom is a transverse slit. The perfect shot acquire enough motion to carry them safely across, while the imperfect ones hobble down and fall into the slot.

135. Further Illustrations of Surface Tension. — A mixture of alcohol and water may be made with the same density as that of olive oil. A small mass of the oil placed in this mixture will assume a globular form, since it is not distorted by its weight. If the attending conditions do not permit it to become spherical, it will in every case assume a form having the smallest surface under the given conditions. If, for example, a metallic ring be immersed in a large globule of olive oil suspended in the mixture and some of the oil be then removed by means of a pipette, the remainder will take the form of a double convex lens.

Fig. 65

Make a stout wire ring three or four inches in diameter with a handle (Fig. 65). Tie to this a loop of thread so that the loop may hang near the middle of the ring. Dip the ring into a good soap solution containing glycerine, and obtain a plane film. The thread will float in it.

Break the film inside the loop with a warm pointed wire, and the loop will spring out into a circle. The tension of the film attached to the thread pulls it out equally in all directions. By tilting the ring from side to side, the circle may be made to float about on the film.

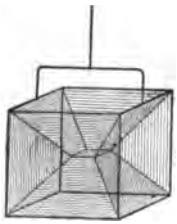


Fig. 66

Interesting surfaces may be obtained by dipping skeleton frames made of stout wire into a soap solution. The films in Figure 66 are all plane, and the angles where three surfaces meet along a line are necessarily 120° for equilibrium.

A small bit of camphor gum placed on warm water, perfectly free from any oily film, will execute rapid gyrations across the surface. The camphor dissolves unequally at different points, and thus produces an unequal weakening of the surface tension in different directions. An interesting modification of this experiment is to make a miniature tin or wooden boat, with a notch in the stern to hold a bit of camphor gum. The camphor weakens the tension astern, and the tension at the bow draws the boat forward.

136. Surface Energy and Surface Tension. — If we call the loss of potential energy, due to a diminution in the surface of one unit, the *surface energy* per unit area, it can be shown that this is numerically equal to the surface tension *per unit width of the film*. The following should be regarded as illustrating this relation rather than a model method of measurement.

Let a soap film be stretched on a frame BCD (Fig. 67) with the light rod A movable. Denote the length of the rod between B and D by a, and let the rod be drawn downward a distance b. Then the increase in the surface for the two sides of the film is $2\,ab$; and if E is the superficial energy per unit area, the increase in the potential energy of the surface is $2\,Eab$.

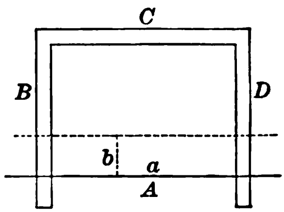


Fig. 67

The work done against the surface tension perpendicular to A may be placed equal to the increase in surface energy.

Let T be the surface tension *per unit width*; then the whole tension in the two sides of the film is $2\,Ta$. The work done against this force through the distance b is $2\,Tab$. Hence the equation

$$2\,Tab = 2\,Eab,$$

or

$$T = E.$$

The surface tension *per unit width* is therefore equal to the surface energy *per unit area*.

For pure water and air the surface tension is 75.8 dynes per centimeter width; for mercury it is 527.2 dynes, both at 0° C.

137. Theory of the Spread of Oil on Water. — Suppose a drop of oil to be placed on water (Fig. 68). There are then three fluids in contact: (a) air, (b) water, and (c) oil, and three surface tensions act on a particle at O; namely, T_{ab}, between air and water, T_{ac}, between air and oil, and T_{bc}, between water and oil, in the direction of the arrows respectively. Then if the three tensions are in equilibrium, they may be represented by the sides of a triangle taken in order (§ 102), and the angles made by the three surfaces will depend only on the relative magnitudes of the three surface tensions.

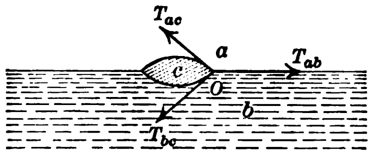


Fig 68

If now $T_{ab} > (T_{ac} + T_{bc})$, the two forces T_{ac} and T_{bc} cannot be in equilibrium with the third force T_{ab}, and the particle at O must move in the direction of the force T_{ab}.

For air, water, and oil the tensions are as follows:

$T_{ab} = 75.8$ dynes; $T_{ac} = 36.88$ dynes; $T_{bc} = 20.56$ dynes, all per centimeter width.

Then $$75.8 > (36.88 + 20.56),$$

and the oil at O moves until the whole surface of the water is covered with a thin film of oil. Although the oil is spread by the pull of the superior surface tension, it is customary to

say that oil spreads itself on water. The so-called oiling of the sea takes place in this manner.

If a small drop of turpentine be placed on clean water by means of a thin glass rod, and the surface be viewed by strong reflected light, the turpentine may be seen spreading over the water like a flash of light.

138. Pressure within a Bubble. If a soap bubble is blown on a large thistle tube and the open end of the tube is held near a candle flame, the contraction of the bubble on account of surface tension will expel the air with sufficient force, it may be, to blow out the flame. The surface tension produces a normal pressure inward in excess of atmospheric pressure. Call this excess p per unit area. It may be expressed in terms of surface tension per unit width and the radius of the bubble as follows:

Imagine the bubble to be divided into two halves by a plane through its center. Then the pressure on the two halves of the bubble normal to this plane, which is the sum of all the components of p normal to the plane on one of the halves, is the area of the section times the pressure p, or $\pi R^2 p$, R being the radius of the bubble.

This pressure is equivalent to the surface tension all around the circumference bounding the section, which for both inside and outside surfaces equals $2\,T \times 2\,\pi R$, or $4\,\pi R T$. Then

$$\pi R^2 p = 4\,\pi R T,$$

and

$$p = \frac{4T}{R}.$$

From this expression it will be seen that the pressure inside a bubble increases as the bubble gets smaller. If fog and cloud consisted of small vesicles of water, as some have supposed, they would still be heavier than the air displaced, both because of the weight of the water and because the air within is under greater pressure than the pressure of the atmosphere on the outside.

139. Capillary Phenomena.—If a fine glass tube, commonly called a capillary tube, be partly immersed vertically in water, the water will rise higher in the tube than the level outside; and the smaller the diameter of the tube, the higher will the water rise (Fig. 69). On the other hand, mercury in the tube is depressed below the level outside.

Similarly, if two chemically clean glass plates, inclined at a very small angle, be supported with their lower edges in water, the height to which the water rises at different points is inversely as the distance between the plates at the points, and the water line is a curve known as a rectangular hyperbola (Fig. 70). By coloring the water slightly this curve may be readily projected on a screen.

Fig. 69

It is easy to determine that the free surface of a liquid is not horizontal near the sides of the vessel containing it, but is noticeably curved. When the liquid wets the vessel, as water in glass, the surface is concave and the water rises along the glass; when the liquid does not adhere, as mercury in glass, the surface is convex and the mercury is depressed.

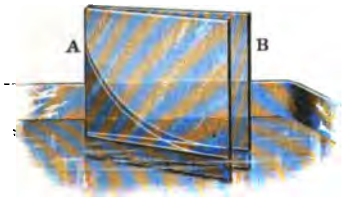


Fig. 70

The elevation or depression varies with the material of the tube and the nature of the liquid. Water in glass rises to a higher level than any other liquid. Capillary elevation for water is nearly three times as great as for sulphuric ether or bisulphide of carbon. A rise of temperature causes a decrease in the elevation and the depression respectively.

Familiar examples of capillary action are numerous. Blotting paper absorbs ink in its fine pores, and oil rises in a wick, by capillary action. A sponge absorbs water for the same reason. The spread of water through a lump of sugar may be similarly explained. Small objects drift together on water or cling to the sides of the vessel because of capillary action; and for the same reason water rises around a fine wire and interferes with its free rotation.

140. Capillary Elevation and Depression Due to Surface Tension. The surface tension between air and a liquid acts around the inner circumference of the tube, downward in the

case of a convex *meniscus*, as the curved surface is called, and upward in the case of a concave meniscus. Let h (Fig. 71) be the mean elevation of the liquid in the tube. Then if the liquid wets the wall of the tube, the angle θ is very nearly zero. The entire surface tension around the interior of the tube where the film adheres is $2\pi r T$, r being the inner radius of the tube, and this force acts vertically. The force, lifting the liquid and depressing the tube, is in equilibrium with the weight of the liquid column of the height h. Let s be the weight of unit volume of the liquid. The weight of the whole column is $\pi r^2 hs$. Consequently,

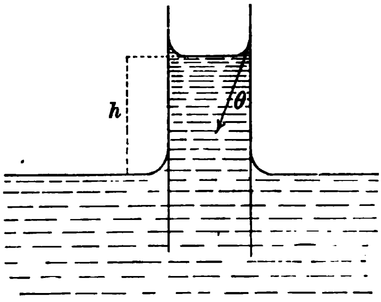


Fig. 71

$$2\pi r T = \pi r^2 h s,$$

$$h = \frac{2T}{rs}, \tag{41}$$

or the elevation is inversely as the radius or the diameter of the tube. This expresses the law of capillary elevation and depression.

141. Osmosis. — If layers of chloroform and sulphuric ether are separated by a layer of water in a closed bottle, and are left undisturbed, in course of time the ether will pass into the chloroform. The ether dissolves to some extent in the water and is removed from the other side by the chloroform. The water does not permit the chloroform to pass to an appreciable extent into the ether.

A thin sheet of India rubber between alcohol and water allows the alcohol to pass into it and to be removed from the other side by the water. This process is properly called diffusion through a membrane.

When a porous partition is employed between the two substances, the passage of two liquids through it at different rates is called *osmosis*. Tie a piece of dampened parchment over the large end of a thistle tube so that it shall be water-tight. Fill to the bottom of the funnel part with a saturated solution of copper sulphate and immerse in water, so that the level of the liquids outside and in shall be the same. The level of the liquid within the tube will rise slowly, while the water outside will acquire a bluish tint; thus, while some copper sulphate passes out through the membrane, more water passes in and raises the level inside. The process continues until the hydrostatic pressure in the long thistle tube produces an equilibrium of flow in the two directions through the membrane.

142. Osmotic Pressure. — Pfeffer devised a porous partition by depositing ferrocyanide of copper in the pores of an unglazed porcelain cylinder. Such a partition is called a semipermeable membrane, because it is permeable to water in one direction and only slightly to a solution of sugar in the other. The cylinder, filled with a solution of cane sugar, was submerged in water, and the pressure inside required to equalize the flow in the two directions was measured by Pfeffer by means of a column of mercury. The pressure is called *osmotic pressure*. Some of his results are the following:

Percentage of Sugar in Solution	Osmotic Pressure p in cm. of Mercury	Mass of Sugar per cu. cm. of the Solution	Volume v containing 1 Gram-molecule (342 gm.) of Sugar	Product pv.
1	53.5	0.01004	34,064	1,822,400
2	101.6	0.02016	16,968	1,723,600
4	208.2	0.04063	8,417	1,752,500
6	307.5	0.06119	5,589	1,718,600

These measurements were made at the same temperature, and it will be observed that the product pv is a fairly constant quantity.

Measurements of the change of osmotic pressure with temperature have shown that for a one per cent sugar solution the following relation holds:

$$p = 49.62(1 + 0.00367\,t).$$

The constant 49.62 is the pressure p_0 at 0°. The coefficient $0.00367 = 1/273$ is the same as the coefficient of the change of pressure for air with change of temperature.

A further discussion of these relations must be reserved for a later chapter in the subject of Heat.

II. PRESSURE IN FLUIDS

143. Laws of Fluid Pressure.—The three fundamental characteristics of pressure in fluids at rest may be called the laws of fluid pressure. They are:

I. *Fluid pressure is normal to any surface on which it acts.*

II. *Fluid pressure at a point in a fluid is of the same intensity in all directions.*

III. *Pressure applied to a fluid is transmitted undiminished in all directions.*

Fluid pressure is measured in terms of the pressure per unit area. By pressure at a point is meant the pressure per unit area at the point. If the pressure over a surface is not uniform, then we may consider only a surface about the point so small that the pressure over it is practically constant, and the quotient of the force on it by this small area will be the pressure per unit area.

144. Pascal's Principle.—The first of the laws of fluid pressure is a consequence of the mobility of a fluid. If the pressure is not perpendicular to the surface, it can be resolved into a normal component and one parallel to the surface. This latter component would produce motion of the fluid parallel to the surface; but since the fluid is assumed to be

at rest, no such motion can take place, and therefore the pressure exerted by a fluid on any surface is normal to that surface at every point.

The other two laws are included in Pascal's principle enunciated in 1653. It is founded on the equal transmission of pressure in all directions. A solid transmits pressure only in the direction in which the force acts; but a fluid, either a liquid or a gas, transmits pressure in all directions. Hence Pascal's fundamental law of the mechanics of fluids:

Pressure applied to an inclosed fluid is transmitted equally in all directions and without diminution to every part of the fluid and of the walls of the containing vessel.

If a small cubical element of the fluid anywhere in the interior be imagined solidified without other change, this element will remain in equilibrium however it be turned about. But for equilibrium the forces acting on all the faces of this small cube balance, or the pressures on all the faces are the same. But the faces all have the same area; therefore the intensity of the pressure on all the faces is the same.


Fig. 72

If a thin-walled bottle be filled with water and a close-grained cork be fitted to it, pressure applied to the cork (Fig. 72), with a lever if necessary, may cause the bottle to break, especially if it has flat sides. The whole bursting force is equal to the product of the area of the inner surface of the bottle and the pressure per unit area. Thus, if the inner surface be 40 square inches and the force applied to the cork 25 pounds per square inch, the bursting pressure is 1000 pounds of force.

When a balloon is inflated by pumping into it illuminating gas or hydrogen, the balloon swells out equally in all directions. This fact of equal pressures in all directions in the gas is evident in the inflation of a toy balloon or a soap bubble.

145. The Hydraulic Press. — Pascal's principle has an important application in the hydraulic press invented by Joseph Bramah in 1795, and hence sometimes called a Bramah press.

It is employed for exerting great pressure, as in baling cotton, making lead pipe, and lifting heavy masses of metal in steel mills, locomotive works, and on warships.

Figure 73 is a section showing the principle. A metal piston passes water-tight through the collar *n* of the large cylinder, while the piston *p* is worked up and down as a force pump, pumps water from a reservoir at the bottom and forces it through the pipe *K* into the cylinder *B*. When the plunger *p* is descending, the water transmits the applied pressure to the base of the large piston or ram, which is thus

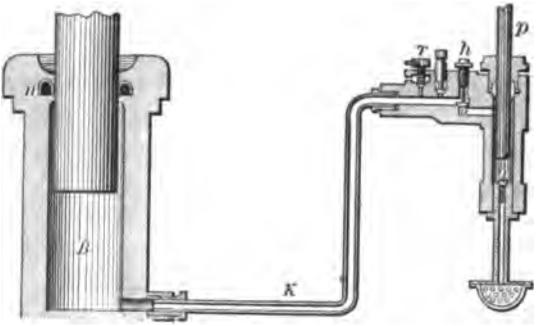


Fig. 73

forced up with its load. If the cross-sectional area of the plunger *p* is *a* units and the force applied to it is *P*, the pressure per unit area on the water is P/a. This pressure is transmitted to *B* and acts on each unit of surface there. If the area of the base of the piston in *B* is *A* units, the force *W* acting on the piston is AP/a. Hence the mechanical advantage is

$$\frac{W}{P} = \frac{A}{a} = \frac{D^2}{d^2},$$

where *D* is the diameter of the ram and *d* that of the pump plunger.

If friction be neglected, this machine conforms to the principle of work, for it is evident that the small piston travels as much farther than the large one as the force exerted on

the large piston is greater than the effort applied to the plunger of the pump.

146. Fluids acted on by Gravity.—The weight of each horizontal layer of a fluid at rest is transmitted to every layer at a lower level. The pressure in the lower layers is then greater than in the upper ones, since each layer supports the weight of all those above it. But the pressure throughout any horizontal layer is everywhere the same; otherwise the fluid would flow from points of higher pressure in the horizontal plane to those of lower pressure, since no work would have to be done against gravity, so long as the motion is in the same horizontal plane.

Fig. 74

The pressure on any horizontal plane in a liquid due to the weight of the liquid itself is proportional to its depth below the surface; for at every point of such a plane is supported a column of the liquid of the same height, and this height is the depth of the layer.

Moreover, the pressure at a point is the same in every direction. If three glass tubes, bent as shown in Figure 74, be filled to the same height with mercury, when they are immersed so that their inner open ends are at the same level, the bottom of the bends resting on the bottom of the tall jar, the difference in level of the mercury is the same in the three tubes; so that at any point the pressure downwards, sideways, and upwards is the same. Also, the pressure measured in this way will be found to be proportional to the depth.

It is immaterial whether the pressure on any horizontal plane is due to the weight of the liquid or is in part due to

an externally applied pressure. The equality of pressure in all directions is a consequence of the equal transmission of pressures in all directions. All the pressures in a liquid at rest must be balanced, since unbalanced pressure would produce currents in the liquid.

147. Pressure on the Bottom of a Vessel Independent of its Shape. — Vessels known as Pascal's vases, made to screw into the same ring base with a removable bottom (Fig. 75), are used to demonstrate that the pressure on the detachable bottom is independent of the shape of the vessel. The bottom is suspended from one arm of a balance with an appropriate weight in the scale pan on the other side. The three vessels are successively screwed into the same base; it is found that the bottom is detached and allows the water to escape when it reaches the same height, notwithstanding the difference in the weight of water required to fill the three vessels to the same level.

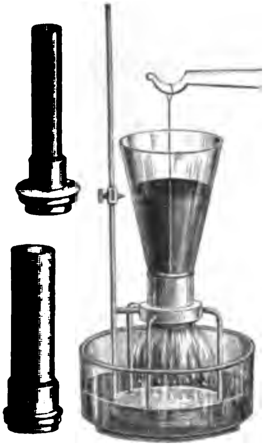


Fig. 75

Therefore, *the pressure on the bottom of a vessel is independent of the shape of the vessel.* The apparent contradiction involved in the fact that unequal masses of water produce equal pressures has often been called the *hydrostatic paradox.*

148. Total Pressure on an Immersed Surface. — Liquid pressure depends on the depth and on the weight of a unit cube of the liquid. The pressure on any horizontal area may be calculated as follows:

Let A denote the area pressed upon, H its depth below the surface, and w the weight of a unit of volume of the liquid. Then the pressure is equal to the weight of a cylindrical

column of the liquid, the base of which has an area A, and its height H. It is therefore

$$P = AHw. \tag{42}$$

In the metric system, w is 1 gm. per cubic centimeter; in the English system, w is 62.4 lb. per cubic foot, both for water.

If the surface pressed upon is plane but not horizontal, then the average pressure over its surface will be the pressure on a unit area at the depth of its center of figure, or center of gravity. For the whole pressure on such a surface the expression is again

$$P = AHw,$$

where H is the depth of the center of gravity of the surface.

Hence the general expression for the pressure on a surface immersed in any liquid is, *the weight of a column of the liquid, the base of which is equal in area to the surface pressed upon, and its height, the depth of its center of gravity below the surface of the liquid.*

149. Surface of a Liquid at Rest.—The free surface of a liquid at rest is horizontal. Consider a particle m at some point B of the free surface ABD (Fig. 76) which we may suppose is not horizontal. The vertical force on m is mg, represented by BW. This force may be resolved into two components, one normal to the surface, and the other, BC, parallel to the surface. Since the air pressure on the surface is everywhere the same, there is no hydrostatic pressure to resist this latter component of pressure; and as there is no friction of rest in a liquid, the particle m at B must move. When the surface is level, BC vanishes and there is no motion.

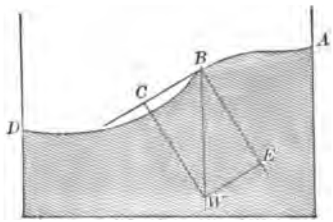


Fig. 76

The sea or any large expanse of water is a part of the

spheroidal surface of the earth. When one looks with a field glass at a long straight stretch of the Suez Canal, the water and the retaining wall, as contrasting surfaces, may be seen to curve over in a vertical plane.

III. BODIES IMMERSED IN LIQUIDS

150. Buoyancy. — An iron ball sinks in water and floats in mercury. An egg sinks in fresh water and floats in a saturated solution of common salt. A piece of oak floats in water, but the dense wood lignum-vitæ sinks. When a swimmer wades up to his neck in sea water, he is nearly lifted off his feet by the heavy salt water which buoys him up.

The resultant of the upward pressure of a liquid on a body immersed in it is a vertical force, and it counterbalances a part or the whole of the body's weight. This resultant upward pressure of a fluid is called *buoyancy*.

Suspend a brass or iron weight from the hook of a spring balance and note the weight. Now bring a beaker of water up under the weight and submerge it. Its apparent weight will be diminished. If salt water is used, the apparent loss of weight will be greater; if kerosene, it will be less. In the shortened form of popular language the body immersed is said to have suffered a loss of weight, though its real weight has not changed in the least. Another force has been brought to bear on it, namely, the excess upward pressure of the liquid, called buoyancy.

151. The Principle of Archimedes. — The law of buoyancy was discovered by Archimedes about 240 B.C. It is as follows:

A body immersed in a liquid is buoyed up by a force equal to the weight of the liquid displaced by it.

Suppose a cube immersed in water (Fig. 77). The pressure on its vertical sides *a* and *b* are equal and in opposite directions. The same is true of the other pair of vertical faces. The resultant horizontal pressure is therefore zero. On *d* there is a downward pressure equal to the weight of the column of water having the face *d* as a base and a height

dn. On the bottom *c* there is an upward pressure equal to the weight of a column of water having a base *c* and a height *cn*. The upward pressure therefore exceeds the downward pressure by the weight of a prism of water, the base of which is the face *c* of the cube, and its height the difference between *dn* and *cn* or *cd*; and this is the weight of the volume of water displaced by the cube.

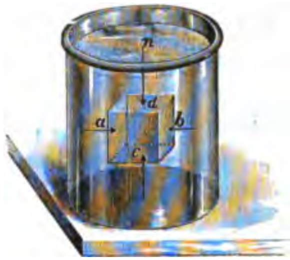


Fig. 77

In general, if we consider any portion of the mass of water solidified without other change, its equilibrium would not be disturbed and its own weight may be considered as acting vertically downward through its center of gravity. The resultant liquid pressure on its surface must therefore be equal to its weight and must act vertically upward through its center of gravity for equilibrium.


Fig. 78

A metallic cylinder 5.1 cm. long and 2.5 cm. in diameter has a volume of almost exactly 25 cm.^{3} Suspend it by a fine thread from one arm of a balance (Fig. 78) and counterpoise. Then submerge it in water as in the figure. The equilibrium will be restored by placing 25 gm. in the pan above the cylinder. The cylinder displaces 25 cm.^{3} of water weighing 25 gm. and its apparent loss of weight is also 25 gm. Strictly the temperature of the water should be 4° C.

152. Equilibrium of Floating Bodies. — A body cannot float partly immersed in a liquid unless it is specifically lighter than the liquid. If no other force acts on the body, it will sink until the weight of the displaced liquid exactly equals that of the body.

In liquids the buoyancy is practically independent of the

depth so long as the body is completely immersed, but it will decrease as soon as the body begins to emerge from the liquid. Hence,

When a body floats on a liquid, it sinks to such a depth that the weight of the liquid displaced equals its own weight.

The weight of a body acts vertically downward, and the resultant pressure of the liquid acts vertically upward through the center of gravity of the displaced liquid, called its *center of buoyancy*, for equilibrium. These two forces must be equal and in the same vertical line.

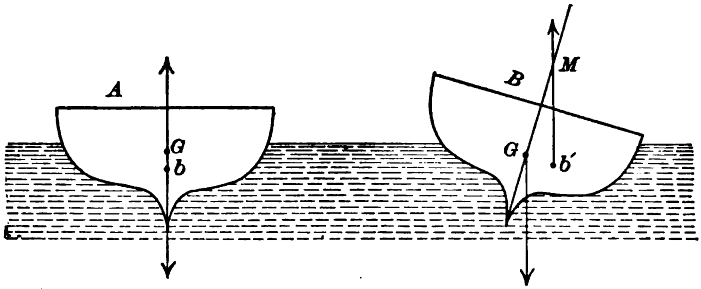


Fig. 79

Let Figure 79 be the section of a floating body, such as a boat, G its center of gravity, and b the center of buoyancy. Then for equilibrium G and b must be in the same vertical line, as in A. If the body be displaced by tilting, as in B, the center of buoyancy will also be displaced to some point b'. The floating body is then acted on by a couple consisting of the upward pressure through b' and the weight of the body acting downward through G. This couple tends to restore the body to its position of equilibrium. The point M, where the new vertical through b' cuts the medial line Gb, is called the *metacenter*. For angular stability the metacenter must be above the center of gravity of the floating body, and the farther above the greater the angular stability. If the center of gravity of the floating body is so high and its shape such that its center of gravity is displaced further in tilting than the center of buoyancy, then the metacenter will be below, G the couple acting on the body increases the displacement, and the body is in unstable equilibrium.

A floating body is stable so long as the metacenter remains above its center of gravity. If the metacenter coincides with the center of gravity, as in the case of a uniform sphere, the body is in neutral equilibrium;

but if the metacenter is below the center of gravity, the equilibrium is unstable. The effect of ballast in a ship is to lower its center of gravity and so to increase its angular stability.

153. The Cartesian Diver. — Descartes illustrated the principle of Archimedes by means of a hydrostatic toy, called the *Cartesian diver*. It is made of glass, is hollow, and has a small opening near the bottom. The figure is partly filled with water so that it just floats in a jar of water (Fig. 80). When pressure is applied to the sheet rubber tied over the top of the jar, it is transmitted to the water, more water enters the floating figure, and the air in it is compressed. The figure then displaces less water and sinks. When the pressure is relieved, the air in the diver expands and forces water out again. The actual displacement of water is then increased and the figure rises to the surface. The water in the diver may be so nicely adjusted that the little figure will sink in cold water, but will rise again when the water has reached the temperature of the room and the air in the figure has expanded.

Fig. 80

A good substitute for the diver is a small homœopathic vial in a flat 12-oz. prescription bottle filled with water and closed with a rubber stopper. By pressing on the flat sides of the bottle, the bottle yields, the air in the diver is compressed, and it sinks.

IV. DENSITY AND SPECIFIC GRAVITY

154. Density. — The *density* of a body is *the number of units of mass of it contained in a unit of volume.* In the *c. g. s.* system it is the number of grams per cubic centimeter. If m denotes mass, v volume, and d density, then

$$d = \frac{m}{v}, \quad v = \frac{m}{d}, \quad \text{and } m = vd. \qquad (43)$$

155. Specific Gravity. — The *specific gravity* of a body is the ratio of the mass of any volume of it to the mass of the same volume of pure water at 4° C. Specific gravity is, therefore, only the relative density as compared with water. The specific gravity of solids and liquids is numerically equal to

their density when the latter is expressed in grams per cubic centimeter, since the density of water is then unity.

Let m be the mass of a body and m' the mass of an equal volume of the standard, as water. Then the specific gravity $s = m/m'$ and $m = m's$. If the mass of the water is expressed in pounds and its volume in cubic feet, $m' = v \times 62.4$, and $m = v \times 62.4 \times s$.

Since the density of water at 4° C. in the *c. g. s.* system is sensibly unity, there is no occasion to use the term *specific gravity*. Whatever system of units is used to determine specific gravity, the result will be numerically equal to the density in the *c. g. s.* system. Conversely, if the density is determined in the *c. g. s.* system, the numeral expressing the result is always the specific gravity.

156. Density of Solids. — A. *Solids heavier than water.* The density of a solid insoluble in water may be found by weighing the body first in air and then suspended in water. Its apparent loss of weight in water is, by the principle of Archimedes, equal to the weight of the water displaced, that is, the weight of a mass of water of the same volume as the immersed solid. Hence the quotient of the weight in air by the loss of weight in water is the specific gravity; it is also the mass in grams per cubic centimeter of the solid, or its density.

If the water is not at the temperature of maximum density, then the value found by the process just described should be multiplied by the density D of the water at the temperature of the observation, and

$$d = sD.$$

B. *Solids lighter than water.* Employ a sinker heavy enough to make the body sink in water. Counterbalance with the body in the scale pan and the sinker suspended from the same pan by a fine thread and immersed wholly in water. Transfer the body from the scale pan to the water and attach it to the sinker. The weight w', which must be added to the

scale pan to restore the equilibrium, is the weight of the water displaced by the body. It is not necessary to know the weight of the sinker. Then if w is the weight of the solid in air, the apparent density in grams per cubic centimeter is w/w'.

If the solid is soluble in water, a liquid of known density, in which the body is not soluble, must be used instead of water. If D is the density of this liquid, then $d = sD$ as above.

157. Density of Liquids.—A. *By the specific gravity bottle.* A specific gravity bottle (Fig. 81) is usually made to hold a definite amount of distilled water at a specified temperature, for example, 25, 50, or 100 gm. To check this capacity, weigh the bottle empty and dry, and weigh again when filled with distilled water at the temperature marked on the bottle. The difference, corrected for the density of water at the given temperature, will give the volume in cubic centimeters. Then the bottle must be weighed again when filled with the liquid, the density of which is to be determined. The weight of the liquid divided by the capacity of the bottle in cubic centimeters will be the density of the liquid at the temperature marked on the bottle. The bottle must be filled at this temperature every time.

Fig. 81

Fig 82

B. *By a glass sinker.* Weigh the glass sinker suspended by a fine platinum wire, first in air and then in water. The apparent loss of weight will be the weight of the water displaced by the sinker. Then weigh it again when suspended in the liquid. The loss of weight will now be the weight of the same volume as before, namely, that of the sinker. Divide the latter loss of weight by the former and the quotient will be the density of the liquid in grams per cubic centimeter.

C. *By the hydrometer.* The common hydrometer is usually made of glass and consists of a cylindrical stem and a bulb weighted with mercury

or fine shot to make it sink to the requisite level (Fig. 82). The stem is graduated so that the depth to which the instrument sinks can be read off, or at least a reading can be taken at the level of the water which may be made to give the density of the liquid in which the hydrometer floats.

The stems of hydrometers are frequently graduated to give directly the density of the liquids in which they are immersed. In this case the length of the divisions on the stem decreases from the top downward.

Hydrometers are sometimes provided with a thermometer in the stem to indicate the temperature of the liquid at the time of taking the density. Hydrometers of the type described are hydrometers of variable immersion as distinguished from those of constant immersion. The former are extensively used in the arts for the approximate testing of the liquids used. They are generally graduated with reference to their specific use. Special names are then applied to them, such as lactometers for testing milk, alcoholmeters for determining the strength of spirits, acidimeters for testing the strength of acids, etc.

158. Two Liquids in Communicating Tubes.—If the heavier liquid, mercury for example, is first poured into a U-shaped tube (Fig. 83), and then the lighter one, water for example, is poured into one limb, the two liquids of unequal densities will rise to different levels above the surface of separation common to the two. If h and h' are the heights of their free surface above the common plane HH', then for equilibrium the pressures of the two columns h and h' on the common plane are the same.

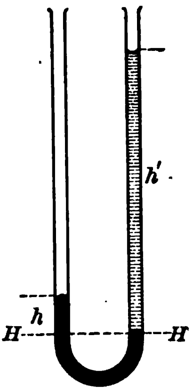


Fig. 83

Let d and d' be the densities of the two liquids. For the pressure on unit area the two columns have the volumes h and h' and masses hd and $h'd'$. Hence their weights are

$$hdg = h'd'g,$$

or

$$\frac{h}{h'} = \frac{d'}{d}.$$

The densities of the two liquids are therefore in the inverse ratio of their heights above the common plane of separation. This relation furnishes an additional method of comparing the densities of two liquids that do not react chemically.

V. FLUIDS IN MOTION

159. Velocity of Efflux. — When a small opening is made in the side of a vessel containing a liquid, at a vertical distance h below the surface, the liquid flows out with a definite velocity v. Torricelli's formula for the velocity of efflux is

$$v^2 = 2gh. \tag{44}$$

This is the same as the velocity acquired by a body falling through a height h in a vacuum.

When a small mass m issues from the orifice, an equal mass m falls some distance a_1 to take its place; another equal mass must in consequence fall a distance a_2, and so on through a series to the surface.

The total loss of potential energy of the liquid is then

$$mga_1 + mga_2 + mga_3 + \cdots = mgh,$$

where h is the sum of the distances a_1, a_2, a_3, etc., or the height of the free surface of the liquid above the orifice.

Then, neglecting viscosity, the loss of potential energy should equal the energy of motion acquired by the mass m. If so, we may write

$$mgh = \tfrac{1}{2} mv^2.$$

Solving for v^2, we have

$$v^2 = 2gh,$$

which is Torricelli's formula.

If the area of the orifice is a, the quantity of liquid discharged in time t should be avt. In point of fact the rate of discharge is less than this. If the opening be a simple orifice in the side of the vessel, the quantity of liquid discharged is about 62 per cent of the quantity avt. The difference is due chiefly to the convergence of the stream lines, which produces a contraction of the jet just outside the orifice. By the use of an orifice or projecting mouth-piece conforming to the conical shape of the contracted stream, a velocity but little short of $\sqrt{2gh}$ may be attained.

160. Range of Jets. — Let ED (Fig. 84) be the vertical side of a vessel of water, the surface being at E. EA, AB, BC, and CD are equal distances. Let v be the velocity of efflux from the orifice at A, the depth of which below the surface is h and its height above the horizontal plane through D is b. Then if t is the time of falling through the height b, the horizontal range of the stream is

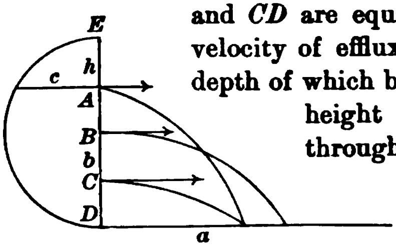


Fig. 84

$$a = vt. \qquad (a)$$

Also from equation (20),

$$b = \tfrac{1}{2}gt^2. \qquad (b)$$

But if $v^2 = 2gh$ by Torricelli's theorem, from (a)

$$t^2 = \frac{a^2}{v^2} = \frac{a^2}{2gh}.$$

Substitute this value of t^2 in (b), and

$$b = \frac{a^2}{4h}.$$

Whence $a^2 = 4bh$ and $a = 2\sqrt{bh}$.

It follows that if b and h exchange values, the range will be the same, for their product bh is not changed. Hence the range from the orifice C is the same as that from A.

The greatest range is from the orifice B midway between the top and the bottom. This may be demonstrated as follows:

On ED as a diameter describe a semicircle. The square of any half chord c is equal to the product of the two segments into which the chord divides the diameter, or

$$c^2 = bh \text{ and } c = \sqrt{bh}.$$

But c is a maximum at the point B. The range $2\sqrt{bh}$ is therefore also a maximum for the orifice B.

161. Flow of Liquids through Tubes.—If a metallic tube be attached to an orifice, the velocity of efflux will be diminished by reason of friction and viscosity. The layer of liquid in contact with the wall of the tube remains nearly at rest, especially if the tube be narrow and rough, and the velocity of flow increases toward the axis. The flow of an open stream of water is greatest near the middle of the stream and at the surface, where the friction is least.

If an elastic rubber tube is attached instead of a rigid one, the efflux will be the same as by a rigid tube of the same length and diameter, so long as there are no sudden variations of pressure. If, however, the pressure be intermittent or pulsating, the stream from a rigid tube reproduces every variation of the pulsating pressure; while an elastic tube rapidly eliminates the inequalities of pressure, and if it be of sufficient length, the pulsations in the stream disappear. The mechanical effect is due to an alternate storage and restoration, or give and take, of energy by means of the elasticity of the tube. This action is analogous to that of a flywheel in producing a steady flow of energy from a reciprocating steam engine. It has its analogue also in alternating currents of electricity. Attention will be drawn to this analogy in a later chapter (§ 651).

162. Flow of Liquids in Vertical Pipes.—In a vertical pipe the rapidly descending liquid breaks into sections fitting the pipe more or less perfectly. They acquire increasing velocity in their descent through an unobstructed pipe, and act as liquid pistons. A partial vacuum is thus produced, and the outer pressure of the atmosphere forces the liquid into the pipe all the more rapidly. This partial vacuum in the long waste pipe from a washbasin or a bathtub accounts for the noisy flow of the water, and it may be sufficient even to withdraw the water from a siphon trap and leave the pipe open to the sewer. It is therefore essential that all such traps leading to a sewer should have a separate vent to the outside air.

163. Flow in Pipes of Variable Section. — The flow of liquids through a pipe of variable cross section presents some interesting features. The variations of pressure in such a pipe may be shown by the height to which the liquid rises in vertical tubes attached to the pipe of variable diameter, as in Figure 85. When the flow is such as to keep the pipe full, the pressure is greatest in the widest parts and least in the narrowest. This apparent anomaly admits of ready explanation.

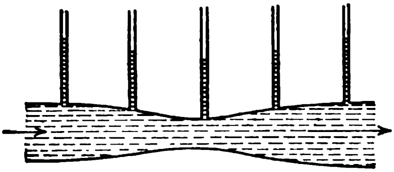


Fig. 85.

It is evident that when the flow is steady, the velocity of the stream is greatest at the narrowest part of the pipe, since the rate of flow past every section must be the same, and the narrower the pipe the greater the velocity for the same rate of flow. Hence, when the liquid passes from a wider to a narrower part of the pipe, it is accelerated. With a pipe of uniform cross section and constant velocity, the loss of pressure is uniform, and is due to liquid friction against the pipe and to viscosity. But to produce an acceleration, additional pressure is required, or the pressure must fall from the wider section toward the narrower, in addition to the fall due to friction and viscosity.

On the other hand, when the liquid flows from a narrow to a wide part, it loses velocity, or suffers a retardation. Hence the pressure ahead must be greater than behind, or the pressure increases from the narrower toward the wider section ahead.

VI. PROPERTIES OF GASES

164. Diffusion. — The process by which two fluids mix independently of external pressure is called *diffusion*. A gas diffuses very freely. Insert a rubber tube into the upper part of an inverted jar, about 30 cm. high, and allow illuminating gas to flow in. It will displace the heavier air and

fill the tall jar. Now place the vertical jar filled with illuminating gas over the open end of another jar of the same size, and containing only air. The edges at the open end should be ground and smeared with grease to make an air-tight joint.

Mixing of the two gases cannot take place by reason of gravity, for the lighter gas is above. Nevertheless, in the course of ten minutes there will be found in the lower jar an explosive mixture; the lighter illuminating gas has diffused downward into the air, and the air has likewise diffused upward into the coal gas.

The process of diffusion may be explained by the kinetic theory of gases, which supposes that the molecules of a gas are incessantly moving and colliding with one another. The mutual encounters between the gas molecules and the air molecules diminish the rapidity with which the two gases intermingle. Moreover, the rapidity of the process depends on the nature of the two diffusing gases, and it is especially rapid when one of them is hydrogen.

165. Dalton's Law of Diffusion. — When the two gases have become uniformly mixed by diffusion, the pressure of the mixture remains the same as that under which the gases were before diffusion, assuming only diffusion and no chemical reaction. Hence Dalton's law, according to which each gas in a mixture exerts the same pressure which it would exert if it were alone present; and the pressure on the walls of the containing vessel is the sum of the partial pressures of the separate gases. If, for example, 21 parts by volume of oxygen and 79 parts of nitrogen are placed in a vessel and diffuse into each other under a pressure p, then after the completion of the diffusion, the pressure p_1 of the oxygen will be $\frac{21}{100}$ of the whole pressure, and that of the nitrogen p_2 will be $\frac{79}{100}$, and

$$p_1 + p_2 = p.$$

This is the mixture composing the major part of our atmosphere.

166. Effusion. — The passage of a gas through the fine pores of a partition is called *effusion*. A Florence flask nearly filled with water is surmounted with a jet tube and a funnel tube, to which is cemented air-tight a small battery jar or other porous pot (Fig. 86). A stream of hydrogen is allowed to flow into the beaker inverted over the porous cup. If all the joints are tight, water will issue from the jet tube as a small fountain. The hydrogen passes through the fine pores of the unglazed cup by effusion.

Fig 86

The rate of effusion of different gases is inversely proportional to the square root of their densities. Hydrogen, for example, passes through a porous wall by diffusion four times as fast as oxygen, the density of which is sixteen times as great as that of hydrogen.

167. Boyle's Law. — The relation between the volume of a gas and the pressure to which it is subjected was discovered by Robert Boyle in 1662. It is therefore known as Boyle's law. In France it is called Mariotte's law, from Mariotte, who announced the law fourteen years later than Boyle. The law is as follows:

At a constant temperature the volume of a given mass of gas varies inversely as the pressure to which it is subjected.

If v and p are corresponding volume and pressure, then when the pressure is changed to p', the volume becomes v', and the relation between volumes and pressures is

$$\frac{v}{v'} = \frac{p'}{p} \text{ or } pv = p'v', \tag{45}$$

that is, $pv = c$ a constant.

Boyle's experiments were made with a U-tube (Fig. 87) and they extended only from $\frac{1}{30}$ of an atmosphere to 4 atmospheres. The short leg A was closed at the top and mercury was introduced until it stood at the same level in both legs. The volume of air imprisoned in the short leg, which was under the same pressure as the atmosphere outside (§ 171), was noted and more mercury was then poured into the tube. The difference in the level of the mercury in the two legs of the tube gave the excess pressure above that of the atmosphere in the open limb. The volume of air was read each time from the shorter leg, which had previously been calibrated.

Fig. 87

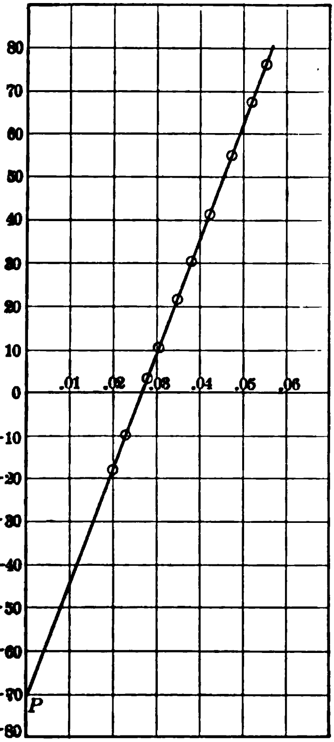


Fig. 88

Equation (45) may be put in the form $p = c\,\frac{1}{v}$, in which c is a constant. This is the equation of a straight line. Hence, if the pressures in terms of the excess column of mercury be plotted as ordinates, and the reciprocals of the corresponding volumes as abscissas, the result should be a straight line (Fig. 88). Moreover, this line should intersect the axis of pressures at a point P below the zero equal to the pressure of the atmosphere, for the pressure of the atmosphere is a part of the

pressure p under which the gas is. The data for Figure 87 were obtained from a piece of apparatus essentially like a U-tube, but better adapted to secure reliable measurements.

168. Researches of Regnault and Amagat. — The application of Boyle's law over a wide range of pressures was not put to an experimental test until nearly 200 years after its dis-

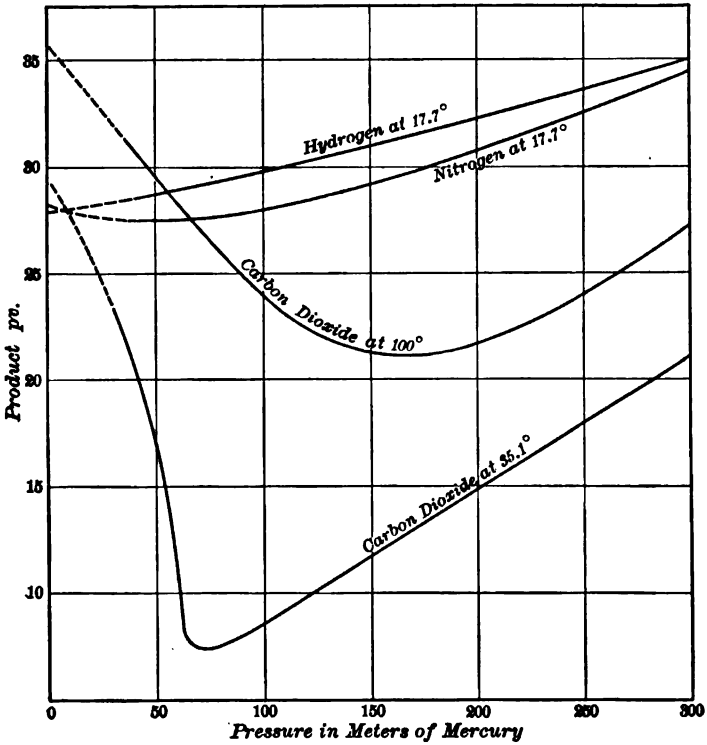


Fig. 89

covery. Regnault carried the pressures as high as 20 meters of mercury. For the first 15 meters he showed that the air and nitrogen are more compressible than they would be if they obeyed Boyle's law with precision, while hydrogen is less compressible.

In 1870 Amagat made use of a coal pit for his experiments on Boyle's law for the reason that it furnished a vertical shaft of nearly constant temperature, and a steel tube 330 m. long could be erected in it. This tube contained the mercury to measure the pressures.

Amagat plotted the products pv as ordinates and the pressures in meters of mercury as abscissas. If Boyle's law were exactly true, the result would be a straight line parallel to the axis of pressures. The curves (Fig. 89) show that hydrogen is from the first less compressible, or more elastic, than the law of Boyle requires; while nitrogen is first more compressible up to a minimum of pv at a pressure of about 40 m. of mercury, and from that pressure upwards is, like hydrogen, less compressible than Boyle's law requires.

The diagram of carbon dioxide shows a more marked departure from a straight line; at 100° C. it is an exaggeration of the nitrogen line, while at 35.1° C. carbon dioxide shows a distinctly marked minimum for pv at a pressure of 70 m. of mercury. At a slightly lower temperature this minimum becomes the line of condensation of the gas into the liquid form.

Gases which at low temperatures and pressures deviate from Boyle's law by being too compressible, at high temperatures and pressures resemble hydrogen and are less compressible, or more elastic, than they would be if they obeyed Boyle's law precisely.

VII. PRESSURE OF THE ATMOSPHERE

169. Air has Weight. — Aristotle attempted to determine whether air has weight by weighing a bladder inflated with air and collapsed. But air, like other fluids, has buoyancy, and the change in buoyancy when the bladder collapsed was equal to the weight of the air removed. Hence Aristotle found no difference in the weight whether the bladder was inflated or not. Since the invention of the air pump it has

been determined that air and hydrogen have the following weights:

1 liter of dry air at 0° C. and under pressure of 76 cm. of mercury weighs 1.293 gm.

1 liter of hydrogen under the same conditions weighs 0.0895 gm.

Therefore the density of air under the above standard conditions is 0.001293 gm. per cubic centimeter; the density of hydrogen, 0.0000895 gm. per cubic centimeter. The density of oxygen is 0.0014279 gm. per cubic centimeter.

170. Torricelli's Experiment.—It was found by Galileo that water would not rise in the pumps of the Duke of Tuscany to a height greater than about 32 feet. He suspected that the pressure of the atmosphere sustained a column of water of this height, but it remained for Torricelli, a pupil of Galileo, to demonstrate in 1643 the truth of Galileo's surmise and to measure the pressure of the atmosphere.

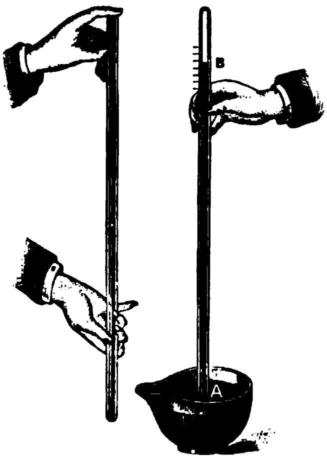


Fig. 90

Torricelli filled with mercury a tube, about a meter long and closed at one end, inverted it, and dipped the open end below the surface of mercury in an open vessel (Fig. 90). The mercury settled down when the finger was removed from the bottom of the tube, and came to a rest at a height of about 76 cm. above the surface of the mercury in the vessel at *A*. Torricelli gave the true explanation to the effect that this column of mercury was supported by the pressure of the atmosphere on the free surface of the mercury in the open vessel. The pressure of 76 cm. of mercury at 0° C. is equiva-

lent to the pressure of 33.9 feet of water at maximum density. Thus Torricelli confirmed the conjecture of Galileo that the water in the pumps of the Duke of Tuscany refused to rise above about 32 feet, not because "nature abhors a vacuum" as the ancients taught, but because the pressure of the atmosphere was insufficient to maintain a higher column.

Torricelli's theory was confirmed by Pascal in Paris, who tried the experiment with tubes filled with oil, water, and wine, and found that the height of the column sustained was inversely proportional to the density of the liquid in the tube. It was Pascal also who suggested that the height of the column of mercury sustained by the pressure of the atmosphere should be less at the top of a mountain than at its base. He verified this prediction by carrying the inverted tube with mercury to the top of the Puy-de-Dôme, about 1000 m. high. A fall in the height of the mercurial column, amounting to about 8 cm., was observed. Thus Torricelli's explanation was completely confirmed.

171. Pressure of the Atmosphere. — The pressure of the atmosphere varies from hour to hour. It is also dependent on the altitude above the sea. It is therefore necessary in defining standard atmospheric pressure, which is measured by the *weight* of a column of mercury of unit cross-sectional area, to define the temperature and the value of the acceleration of gravity. The standard height chosen is 76 cm. of mercury, at a temperature of melting ice (0° C.), and at sea level in latitude 45°. The density of mercury at 0° C. is 13.596, and the value of g at sea level in latitude 45° is 980.6 cm. per second per second. Hence standard atmospheric pressure is

$$76 \times 13.596 \times 980.6 = 1{,}013{,}250 \text{ dynes per sq. cm.}$$

This is a little over 10^6 dynes, or a megadyne.

172. The Barometer. — The *barometer* is an instrument based on the Torricelli experiment, and designed to measure the pressure of the atmosphere. It has become a highly important auxiliary in many branches of physics and in meteorology for weather predictions.

Barometers are either (1) liquid barometers, which measure atmospheric pressure in terms of the height of a column of liquid; or (2) aneroid barometers, in which the pressure is measured by the deformation of the thin corrugated cover of an air-tight metal box.

Fig. 91

Mercury is practically the only liquid used in barometers of the first type, since it does not absorb moisture from the air as glycerine, for example, does, and its density is so great that the column sustained by atmospheric pressure is short enough to be manageable.

The simplest form of mercurial barometer is a Torricellian tube in the form of an inverted siphon (Fig. 91). The short arm has a small hole near the top for the admission of air. The height of the mercurial column is the difference between the readings on the two scales at the right; for example, if the upper pointer stands at 78.4 cm. and the lower one at 4.2 cm., the height of the barometer is 74.2 cm. Corrections must be made for temperature and elevation above sea level. A good barometer must contain pure mercury and the mercury must be boiled in the glass tube to expel all air and moisture.

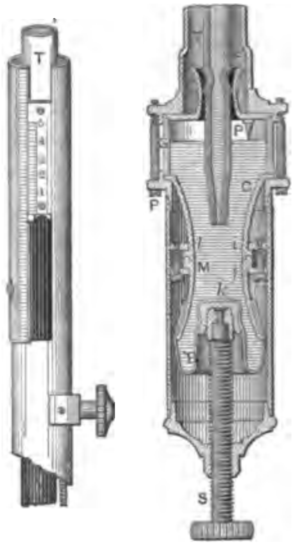


Fig. 92

173. Fortin's Barometer. — A portable barometer designed by Fortin is shown in Figure 92. The cistern C of mercury is closed at the bottom by a leather bag B, the bottom of which may be lowered or raised by means of the screw S. A glass cylinder G near

the top of the cistern permits the observer to see when the ivory point P and its image in the mercury touch each other. The point P is the zero of the barometer scale. The tube is attached to the cistern by means of a piece of chamois skin F, which prevents the escape of mercury but permits no difference of pressure to exist between the outside and the inside air. The glass tube is inclosed in a metal case, which carries a scale and at the top a vernier for measuring fractions of a scale division.

When the barometer is to be transported, the screw S is turned in till the cistern and the tube are entirely filled with mercury. The surging of mercury and the entrance of air are thus prevented.

174. The Aneroid Barometer. — The aneroid barometer consists essentially of a shallow cylindrical box B (Fig. 93), largely exhausted of air; it has a thin cover corrugated in circular ridges to give it greater flexibility. The cover is prevented from collapsing under atmospheric pressure by a stiff spring attached to the center of the cover at M. This flexible cover rises and falls as the pressure of the atmosphere varies and its motion is transmitted to the pointer by means of levers and a chain. A scale graduated by comparison with a mercurial barometer is placed below the pointer. The advantages of an aneroid barometer are its portability and sensitiveness. It should be compared frequently with a mercurial barometer.


Fig. 93

175. Barometric Variations.—Since the mercury in the barometer tube is sustained by the pressure of the air on the mercury outside, changes in the barometric readings indicate corresponding fluctuations in atmospheric pressure. The greater changes follow no well-defined laws, but they herald important atmospheric movements associated with storms.

Experience has shown that rapid barometric changes foretell changes in the weather. Thus a rapid fall of the barom-

eter denotes the near approach of a storm, and a rising barometer is usually followed by fair weather.

The words Rain, Fair, etc., often marked on aneroid barometers especially, are without significance in connection with the readings against which they are placed. Changes in the weather are indicated by rather rapid changes in the readings of the barometer.

176. The Manometer. — A manometer is an instrument designed to measure the pressure of a gas in a closed vessel. A bent tube partly filled with mercury may be used for the purpose. If the pressure to be measured does not greatly exceed one atmosphere, the outer end of the tube is left open, as at *C* in Figure 94 *A*. The end *D* communicates with the closed vessel. The required pressure will then be the reading of the barometer increased by the difference in level between *M* and *L*.

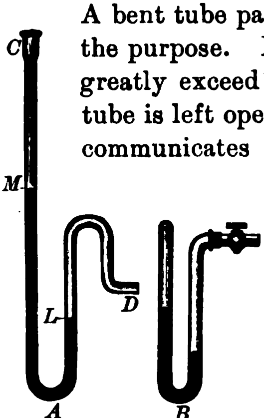


Fig. 94

A similar arrangement may serve for a vacuum gauge. For this purpose, the end *C* communicates with the vacuum apparatus, and the other end is open to the air. The difference in level of the mercury in the two legs of the tube will then be the vacuum.

If the pressure to be measured is considerably greater than one atmosphere, use is made of the elasticity of the air. The outer end of the tube is then closed, as in Figure 94 *B*. The pressure is determined by applying Boyle's law and adding to the result the difference in level of the mercury in the two limbs of the manometer. In practice the pressure is read from a scale attached to the tube and graduated empirically.

In the mechanic arts metallic pressure gauges, analogous to the aneroid barometer, are employed both for pressure and vacuum. They are provided with empirically graduated scales.

VIII. INSTRUMENTS DEPENDING ON PRESSURE OF THE AIR

177. The Air Pump. — The mechanical air pump for exhausting a closed vessel was invented by Otto von Guericke about 1650. The appearance of one of the best forms for general illustrative purposes is shown in Figure 95. The essential internal working parts appear in Figure 96. In the piston *P* is a valve *S* worked mechanically by the motion of the piston rod. The pump cylinder communicates with the outer air at its upper end by a valve *V* working by air pressure, and with the pump plate by a mechanically operated valve *S'*, worked by a rod which passes rather snugly through the piston.

Fig. 95

It is lifted as the upstroke begins, but its ascent is arrested by a stop near the upper end of the rod. During the upstroke the air flows into the cylinder through *S'* from the receiver on the pump plate. When the piston reaches the top of the cylinder, it hits the lever shown in Figure 96, and automatically closes the valve *S'* before the descent begins. In the downstroke the valve *S* opens automati-

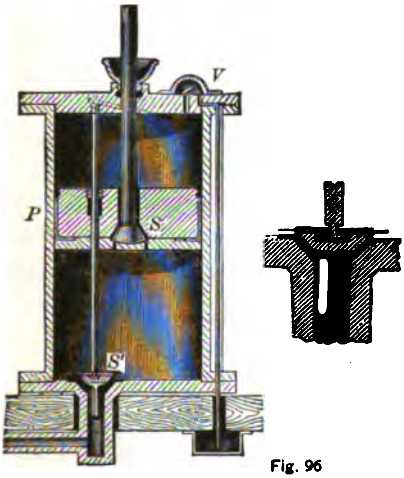


Fig. 96

cally, and the inclosed air passes through it into the upper part of the cylinder. The ascent of the piston again closes *S* and compresses the air above it. If its pressure at the top of the stroke exceeds that of the air outside, the valve *V* opens and air is expelled through the tube leading to the bottom of the cylinder.

Each complete stroke of the pump removes a cylinder full of air; but as the air becomes rarer with each stroke, the mass removed each time is less. On account of leakage, untraversed space, and absorption of air by the lubricating oil, the pressure in the vessel to be exhausted cannot be reduced much below one millimeter of mercury.

178. The Fleuss Pump. — The Fleuss pump (made under the commercial name of "Geryk" pump) is a type of air pump in which a non-volatile oil is used to cover both the piston and the outlet valve. The essential parts are shown in Figure 97. A piston *N* works in a cylinder *M* and around the piston rod is a valve *G* opening outwards. On its upstroke the piston, after passing the inlet *B* from the surrounding chamber, compresses the air in the cylinder and insures the opening of the valve *G* by means of the collar *O*. The oil following leaves no untraversed space. On the downstroke the oil closes the valve *G* completely, a vacuum is produced above the piston, and when the piston reaches the bottom of the cylinder, air flows from any vessel connected with *A* into the pump cylinder.

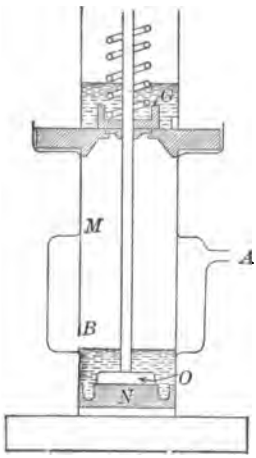


Fig. 97

Two such pumps are often connected in series so that the second one pumps air from the barrel of the first. By such a pump the air pressure can be reduced to a small fraction of a millimeter of mercury; the vacuum produced is sufficient for incandescent lamps and Geissler tubes.

These pumps should always exhaust through a drying tube to prevent the admission of moisture, which is absorbed by the oil and given up again at low pressures. The vacuum will then be reduced to the vapor pressure of water instead of that of the heavy oil used to seal the valves.

179. The Air Compressor. — If the discharge pipe of an air pump were connected to a suitable vessel, while its inlet pipe were left open to the air, air would be forced into the vessel during the action of the pump. Such an arrangement would be an *air compressor*. For a pressure of several atmospheres, valves adapted to withstand high pressures must be employed; and a pump designed for such pressures is suitable for compressing a gas (Fig. 98). The piston is solid and there are two metal valves at the bottom. Air or other gas is admitted through the left-hand tube when the piston rises; when it descends, it compresses the inclosed air, the pressure closes the left-hand valve and opens the other one, and the compressed air is discharged into the compression tank or other receptacle.

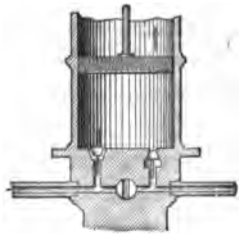


Fig. 98

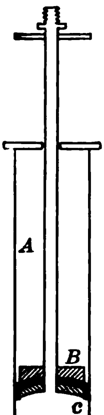


Fig. 99

A bicycle pump (Fig. 99) is an air compressor of a very simple type. In the cylinder A slides a piston B provided with a cup-shaped leather collar C. When the cylinder is pulled outwards, air is admitted past the leather collar, but when the stroke is reversed, a quick motion causes the compressed air to press the leather against the wall of the cylinder and closes it air-tight. The collar thus serves as a valve allowing the air to flow one way but not the other. The compressed air escapes through the tube forming the piston rod. The check valve to prevent the return of the compressed air is the valve in the bicycle tire inlet.

180. Applications. — Both the air pump and the compressor are used extensively in the arts. Sugar refiners employ the air pump to reduce the boiling point of the syrup (§ 405); manufacturers of soda water use a compressor to charge the water with carbon dioxide; in

pneumatic dispatch tubes, now extensively employed for rapidly transporting small packages, both pumps are employed, the one to exhaust the air from the tube in front of the closely fitting carriage, and the other to force compressed air into the tube behind it, so as to propel it with great velocity. Compressed air is also used to facilitate the ventilation of buildings and mines, to operate pneumatic clocks, to control heat regulators, to work air brakes on cars, and to operate pneumatic machinery especially in riveting, calking, and rock boring in tunneling.

181. Buoyancy of the Air.—The principle of Archimedes applies to gases as well as to liquids. The resultant pressure of the atmosphere on bodies is an upward force equal to the weight of the air displaced. A body weighs less in the air than in a vacuum if the volume of air displaced by it is greater than that displaced by the weights.

A *baroscope* is a device to exhibit the upward pressure of the air. A thin hollow globe is slightly overbalanced by a brass weight on a short beam balance (Fig. 100). When the baroscope is placed under a large receiver and the air is exhausted, the hollow sphere sinks, showing that it is really heavier than the counterpoise. In the air it is buoyed up more because its volume is greater than that of the weight. It is perhaps needless to say that the globe must be air-tight.

Fig. 100

182. Balloons.—A soap bubble and a toy balloon filled with air sink because they are heavier than the air displaced; but a bubble filled with hydrogen rises in the air. Its buoyancy is greater than its weight including the hydrogen. The weight of the balloon with its car and contents must be less than that of the air displaced. It is not entirely filled with gas, but as it rises it expands as the pressure of the air outside decreases. Its buoyancy then decreases but little when it rises into a rarer atmosphere.

With hydrogen the upward force is about one kilogram

per cubic meter of gas; with common illuminating gas it is about one half as much. The first ascent in a balloon filled with hydrogen was made by Charles, a professor of physics in Paris, in December, 1783. Gay-Lussac in 1804 ascended to a height of 23,000 feet. At this elevation the barometer sank to 12.6 inches.

The most remarkable ascent ever made was by Messrs. Glaisher and Coxwell in England in 1861. At a height of 29,000 feet, with the thermometer at $-16°$ C., Mr. Glaisher could no longer observe and fainted. According to an approximate estimate, the two men reached an altitude of about 36,000 feet, where the barometer stood at only 7 inches.

In 1900 two long-distance balloon races were made from Paris in an easterly direction. One of the contestants, Count de la Vaulx, the winner in both races, reached Russian territory in both, having traveled the first time a distance of 766 miles in 21 hours and 31 minutes; and the second time, a distance of 1193 miles in 35 hours and 45 minutes. The greatest altitude reached was 18,700 feet.

The aeronauts testify that when the sun shone on the balloon and heated it, the expansion of the gas increased the buoyancy, so that the balloon shot up to higher altitudes. It became necessary in consequence to let out some gas to cause the balloon to descend again. In the night, when the temperature fell, the buoyancy decreased. Ballast was then thrown out to lighten the balloon to prevent its descent. These alternate losses of gas and ballast at length exhausted the capacity of the balloon to keep afloat, and it descended to the ground.

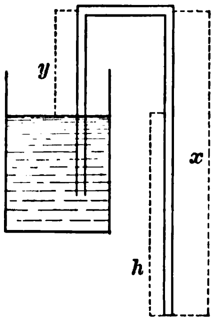


Fig. 101

183. The Siphon. — In its simplest form the siphon is a U-tube employed to convey liquids from one vessel to another at a lower level by means of atmospheric pressure.

Let Y (Fig. 101) be the height of the highest point of the siphon above the liquid in the vessel from which the discharge takes place; and let x be the height of the same point of the siphon above its open end or the surface of the lower

liquid, if the longer arm dips below it. Let H be the height of the column of the liquid equal to atmospheric pressure. Then the pressure outward at the top of the siphon is $(H-y)d$, and the pressure inward at the same point is $(H-x)d$, d being the density of the liquid. The resultant pressure outward is the difference,

$$(H-y)d-(H-x)d=(x-y)d=dh.$$

In the case of water d is unity, and the "head" producing the flow is h. If y exceeds H, the liquid will not rise to the highest point of the siphon by atmospheric pressure and there will be no flow.

A siphon made in the form shown in Figure 102 is called a "vacuum siphon." The short arm ends in a jet tube inside a closed vessel. The pressure within the vessel is less than atmospheric pressure by the weight of a column of the liquid of unit cross section and of a length equal to the vertical distance of the outer end of the discharge arm below the surface of the upper liquid.

Fig. 102

The water in an S-trap may be siphoned off when the discharge pipe is filled with water for a short distance below the trap, unless the trap is ventilated.

184. The Lift Pump. — In the lift pump a piston c, in which there is a valve v', opening upward (Fig. 103), moves nearly air-tight in a cylinder. At the bottom of the cylinder is an opening provided with a valve v, also opening upward. A pipe s leads from this opening down below the surface of the water. When the piston is drawn upward, a partial vacuum is

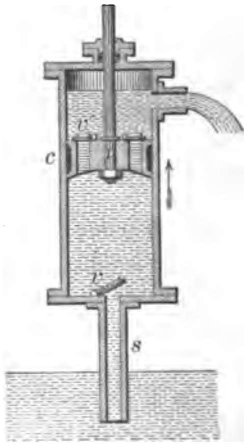


Fig. 103

produced below it, and the pressure of the atmosphere on the water below forces water up the pipe *s* to a height to produce equilibrium. When the piston descends, the valve *v'* opens and *v* closes. Either air or water passes through the upper valve. Finally the upstroke lifts the water above the piston and the pressure of the air on the open water keeps the tube and the piston full. If the piston at the top of its course is less than about 33 feet above the water into which the pipe *s* dips, water will follow the piston to its highest point. The spout may be any reasonable distance above the upper valve; the water above the piston is lifted mechanically and not by atmospheric pressure.

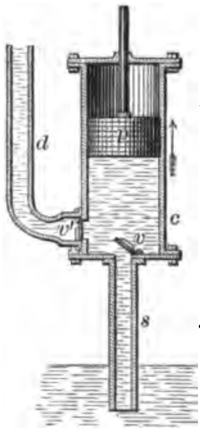


Fig. 104

185. The Force Pump. — The piston of a force pump, like that of an air compressor, is solid and the valve through which the water is forced is below it. Otherwise a force pump is similar to a lifting pump. In Figure 104 *d* is the discharge pipe, *v'* the discharge valve opening outward from the cylinder, and *v* the inlet valve closing the pipe *s*.

In powerful pumps the pipe *d* is surmounted with an air chamber called the *air dome*. The air in it is compressed by water pressure, and it acts as an air spring to give steadiness to the flow of water from the delivery pipe. Fire engines and pumps operated by steam are fitted with an air dome.

Problems

1. The diameters of the cylinders of a hydraulic press are 6 in. and 1 in. respectively. What is the force on the larger piston when a force of 200 lb. is applied to the smaller piston?

2. A hydraulic lift carries a weight of 3000 lb. If the piston supporting the lift is 8 in. in diameter, what pressure of water per square inch will be necessary?

3. A tank 5 ft. deep and 10 ft. square is filled with water. What is the pressure on the bottom? What is it on one side?

4. A cubical block 10 cm. on each edge is submerged in water with its top face 100 cm. below the surface. Calculate the total pressure on the six faces.

5. A vertical tube 2 cm. in diameter is filled with mercury (density 13.6) to a depth of 2 m. What is the pressure per square centimeter at the bottom?

6. What is the pressure per square foot at a depth of 3 mi. in the ocean, sea water being $2\frac{1}{2}$ per cent heavier than fresh water? What is the buoyancy per cubic foot?

7. A body weighs 100 gm. in the air and 88 gm. in water at 20°. Find its density.

8. A glass stopper weighs 150 gm. in the air, 90 gm. in the water, and 42 gm. in sulphuric acid. Calculate the density of the acid.

9. A piece of zinc weighs 70 gm. in air and 60 gm. in water. What will it weigh in alcohol of density 0.8 gm. per cubic centimeter?

10. If the density of sea water is 1.025 gm. per cubic centimeter, and that of ice 0.9 gm. per cubic centimeter, what fraction of an iceberg floating in the sea is under water?

11. A hollow brass ball weighs 1 kgm. What must be its volume so that it will just float in water?

12. A glass tube 72 cm. long and closed at one end is sunk in the ocean with its open end down. When drawn up it was found that the air in the tube had been compressed to within 6 cm. of the top. Assuming a normal atmospheric pressure of 76 cm. of mercury, the density of mercury 13.6 and of sea water 1.025 gm. per cubic centimeter, to what depth did the tube descend?

13. Two bodies, having densities of 6 and 8 gm. per cubic centimeter, respectively, are of such relative volumes that their apparent weights in water are the same. Compare their weights in air.

14. The mark to which a certain hydrometer weighing 90 gm. sinks in alcohol is noted. To make it sink to the same mark in water, it must be loaded with 22 gm. What is the density of the alcohol?

15. A block of wrought iron 10 cm. thick is floated on mercury. To what depth above the mercury must the vessel be filled with water so that the latter shall just reach the top of the block of iron? Densities of iron and of mercury, 7.7 and 13.6 gm. per cubic centimeter, respectively.

16. A liter flask weighing 75 gm. is half filled with water and half with glycerine. The flask and liquids weigh 1205 gm. What is the density of the glycerine?

17. A body floats half submerged in alcohol of density 0.818 gm. per cubic centimeter. What part of its volume would be submerged in water?

18. If the surface tension of a soap solution is 25 dynes/cm., how much greater is the pressure of the gas inside a soap bubble 5 cm. in diameter than that of the air on the outside?

19. If a liter of air weighs 1.29 gm. when the barometer reading is 76 cm., calculate the buoyancy for a ball 20 cm. in diameter when the barometer stands at 70 cm., the temperature being the same.

20. If an open vessel contains 250 gm. of air when the barometric pressure is 76 cm., how much will it contain at the same temperature when the barometric pressure is 70 cm.?

21. The volume of hydrogen gas in a graduated cylinder over mercury was 50 $cm.^3$, the mercury standing 15 cm. high in the cylinder and the barometer reading 75 cm. What would be the volume of the gas if it were under normal pressure?

22. When the barometer reading is 73 cm., what is the greatest possible length for the short arm of a siphon when used for sulphuric acid, density 1.84 gm. per cubic centimeter?

23. A vertical cylinder is closed with a piston whose area is 60 sq. cm. The inclosed air column is 50 cm. high and is at the atmospheric pressure of 1000 gm. per sq. cm. If a weight of 100 kgm. be placed on the piston, how far will it descend, neglecting friction?

24. What is the total pressure on the vertical sides of a cylindrical tank 60 cm. in diameter and filled with water to the height of 2 m.?

25. What is the vertical height of a column of water which will counterbalance a column of benzine 80 cm. high, density 0.9 gm. per cubic centimeter (Fig. 83)?

SOUND

CHAPTER V

WAVES

I. WAVE MOTION

186. Vibrations. — Suspend a heavy ball by a long thread and set it swinging like a pendulum bob. The ball will return at regular intervals to the starting point. If it be set moving in a circle, the thread describing a conical surface, it will again return *periodically* to the point of departure.

A *vibrating* or *oscillating* body is one which repeats its limited motion at regular intervals of time. The term *oscillation* is usually applied to motions of short period only.

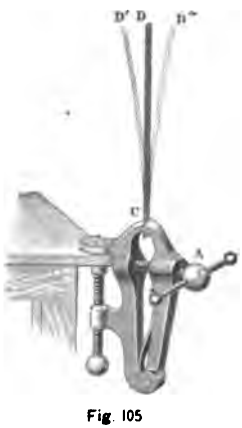


Fig. 105

A *complete* or *double vibration* is the motion between two successive passages of the moving body through any point of its path *in the same direction*, and its *period* is the interval of time taken to execute a complete vibration (§ 92).

Considered with respect to the vibrating body, vibrations are *transverse* when the motion is at right angles to the body. For example, clamp a thin steel strip in a vise (Fig. 105); draw the free end aside and release it. It will move from D' to D'' and back again and its vibrations are transverse.

An example of visible *longitudinal* vibration is the following: Fasten the ends of a long spiral spring securely to

a fixed support with the spring slightly stretched. Crowd together a few turns of the spiral at one end and then release them. A vibratory movement will be transmitted from one end of the spiral to the other, and each turn will swing to and fro in the direction of the length of the spiral (Fig. 106).

Fig. 106

187. Wave Defined. — The periodic motion of a single particle or rigid body has already been studied in Chapter II; we have now to consider the related motions when the various particles of a medium are executing periodic vibrations simultaneously, while the phase of the motion (§ 36) varies from particle to particle in a regular methodic way.

To illustrate: Tie one end of a soft cotton rope to a fixed support. Grasp the other end and stretch the rope horizontally. Start a disturbance by an up-and-down motion of the hand. Each point of the rope will vibrate transversely with simple harmonic motion (§ 36), while the disturbance will travel along the rope toward the fixed end as a wave. This progressive form, due to the periodic vibration of the particles of the medium through which it moves, is called a *wave*. The phase of each successive particle differs from that of the preceding particle by a definite fraction of a whole period.

188. Transverse Waves. — Suppose a series of particles, originally equidistant in a horizontal straight line, to have imparted to them transverse displacements, and that they vibrate with simple harmonic motion. Let the curve (Fig. 107) represent the positions of these particles at some particular instant. They will outline a transverse wave. At g the particle has reached its extreme displacement in the positive direction and is momentarily at rest; the particle at s has reached its maximum negative displacement, and is also at rest. The particle at m is moving in the positive direction with maximum velocity, and the particle at y with maximum velocity in the negative direction. If the wave is traveling toward the right, then an instant later the dis-

placement of g will have diminished, and that of i will have increased to a maximum, the crest having moved forward from g to i in the short interval. The successive particles of the wave all differ in phase by the same amount.

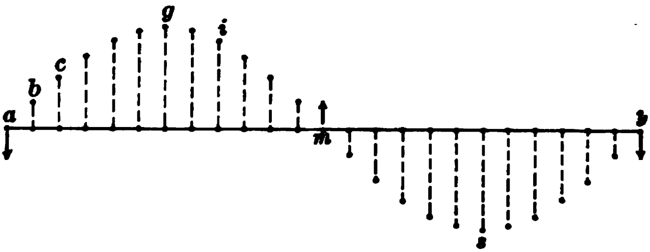


Fig. 107

A *transverse wave* is one in which the vibration of the particles in the wave is at right angles to the direction in which the wave is traveling.

189. Longitudinal Waves.—Place a lighted candle at the conical end of a long tube (Fig. 108). Over the other end tie a piece of parchment or parchment paper. Tap the paper lightly with a cork or rubber mallet; the transmitted impulse will cause the flame to duck, and it may easily be extinguished by a sharper blow. The air in the tube is agitated by the vibratory motion passing through it. A

Fig. 108

wave consisting of a compression followed by a rarefaction traverses the tube, and the dipping of the candle flame indicates the arrival of the compression. The first movement of the membrane compresses the air next to it; the elasticity of the air forces these particles apart again, and this action

compresses the air further along in the tube. The process is continuous and carries the compression through the tube to the flame. Each particle vibrates longitudinally with approximately simple harmonic motion, the whole phenomenon being quite similar to that of the vibrating spiral.

Figure 109 illustrates the distribution of the air particles when disturbed by such a longitudinal wave consisting of

Fig. 109

compressions and rarefactions. *A*, *C*, *E*, etc., are regions of rarefaction; *B*, *D*, *F*, etc., those of compression.

A *longitudinal wave* is one in which the oscillations of the particles composing the wave are to and fro in the same direction as the wave is traveling.

190. Wave Length. — The *length* of a wave is the distance from any particle in the wave to the next one in the same vibration stage, that is, in the *same phase*. Such, for example, is the distance from *a* to *y* in Figure 107, or from *A* to *C*, or *B* to *D* in Figure 109. In a longitudinal wave, a wave length comprises one compression and one rarefaction.

Let λ represent a wave length, n the frequency or number of vibrations a second, and v the velocity of the disturbance or wave. Each particle executes a complete vibration while the wave travels forward a wave length λ. Hence there are n such waves in the distance v. Also, the vibration frequency n and the period T are the reciprocals of each other. Therefore,

$$v = n\lambda = \frac{\lambda}{T}, \text{ or } \lambda = \frac{v}{n}. \qquad (46)$$

II. WATER WAVES

191. Gravitational Waves and Ripples. Waves on the surface of water or other liquid are of a distinct type. For small amplitudes water waves are similar to those arising from

transverse vibrations; but water is so slightly compressible that these waves cannot be waves of compression and rarefaction in a vertical plane. The force by which long waves are produced on the surface of a liquid is the force of gravity. They are therefore commonly called *gravitational waves*.

Suppose by some means the liquid is heaped up in the form *BCD* above the general level and scooped out into a trough *DEF* (Fig. 110). Then the liquid elevated above the general level will tend to flow back toward the level surface, and the upward hydrostatic pressure on the bottom of the trough will cause it to move upward. The simultaneous downward flow of the portion *BCD* into the trough *DEF*, and the rise of the bottom of the trough by upward pressure, cause the forward movement of both crest and trough. Large waves on the surface of the sea are gravitational waves of this character.

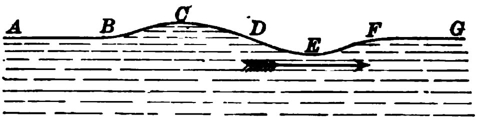


Fig. 110

But gravity is not the only force tending to bring the disturbed surface of a liquid back to a position of stable equilibrium. The surface tension of a liquid (§ 133), which acts like an elastic membrane stretched over the surface, also tends to remove all curvature in the surface, and thus acts as a force of restitution on the displaced liquid particles. The surface tension in a curved surface produces a normal pressure toward the concave side, and this normal pressure increases as the radius of curvature decreases (§ 138). Since the magnitude of the surface tension is small, it comes into account in producing water waves only when the curvature is great, that is, when the waves are short. With short waves the weight of water displaced is small in comparison with the normal force due to surface tension. On the other hand, with long waves of slight curvature, surface tension is negligible in comparison with the gravitational effect. For waves

shorter than 3 mm. or about 0.12 inch, surface tension plays so important a part in the propagation of the wave that the gravity effect is negligible. Waves like these, in which the force of restitution is essentially due to surface tension, are called capillary waves or *ripples.*

For waves longer than 10 cm. or 4 in., surface tension is negligible in comparison with the gravitational force of restitution. The speed of ripples increases as the wave length λ *diminishes;* the speed of gravitational waves increases as the wave length *increases*, provided always that the depth of the liquid is not less than the wave length. There must therefore be a certain wave length for which the speed is a minimum. For water the minimum speed is about 23 cm. or 9 inches per second, and the corresponding wave length is 1.72 cm.

192. Oscillations in Water Waves. — For waves on the surface of deep water the particles describe vertical circles, all in the same plane, containing the direction in which the disturbance is traveling, as illustrated in Figure 111. The circles in the

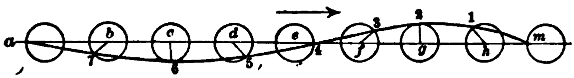


Fig. 111

diagram are divided into eight equal arcs, and the water particles are supposed to describe these circles in the direction of watch hands, all at the same rate; but in any two consecutive circles their phase of motion differs by one eighth of a period. When *a* has completed one revolution, *b* is one eighth of a revolution behind, *c* two eighths or one quarter, etc. A smooth curve drawn through the several simultaneous positions is the outline or contour of a wave.

When a particle is at the crest of a wave, it is moving in the same direction as the wave; but when it is in the trough, its motion is opposite to that of the wave.

As the depth increases, the particles there still move in

circles, but the circles are smaller and smaller as the depth increases, till at the depth of a wave length the radius of the circles is only about $\frac{1}{500}$ as great as at the surface. In shallow water these circular paths become ellipses with their major axes horizontal. The vertical axes decrease with the depth.

The crests and troughs are not of the same size, and the larger the circles (or amplitude), the smaller are the crests in comparison with the troughs. Hence the tops of high waves tend to become sharp or looped, and the waves then break into foam or white caps.

When a high wave comes into shallow water, its speed is diminished; but since the frequency n remains the same, the wave length decreases. The effect of this shortening of the wave length is to increase the amplitude. Such a modification proceeds until an unstable condition is reached, and the wave breaks.

III. PROPAGATION AND REFLECTION OF WAVES

193. Wave Front and Rays. — Suppose that a disturbance originates at a point P (Fig. 112) so that a wave is propagated outwards with the same speed in all directions. After a short interval of time, this disturbance will have spread so as to affect similarly a series of particles at a distance $s = vt$. If the wave spreads in two directions only, the particles at the distance vt will all lie on the circumference of a circle drawn about P with a radius vt. If it extends outwards in three dimensions, the similarly affected particles will lie on the surface of a sphere drawn about the same point as a center and with the same radius. The circumference of the circle in the one case and the surface of the sphere in the other everywhere pass through portions of the medium in the same phase of vibration, and the curve or the surface is called a *wave front*.

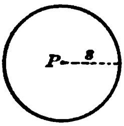


Fig. 112

If the wave front is a sphere, the wave is a spherical wave; if the wave front is either a straight line or a plane, the wave is called a plane wave.

In a uniform medium the direction in which the disturbance is traveling at any point of the wave is normal to the wave front; in the case of a circular or a spherical wave front, this direction is radial. A line drawn to indicate the direction in which the disturbance is propagated, and along which the energy travels outward, is called a *ray*. In an isotropic medium, that is, a medium having the same physical properties in all directions, the wave front and the ray at any point are at right angles to each other. The rays are then straight lines. If the medium is not isotropic, the rays may be curved.

A ray is quite as real as a wave front, and the geometrical methods of treatment founded on rays are as legitimate as the methods depending on wave fronts. In the elementary treatment of subjects involving wave motion it is advisable to make use of either the geometrical or the wave method according as the one or the other may best serve for the simple exposition of fundamental principles.

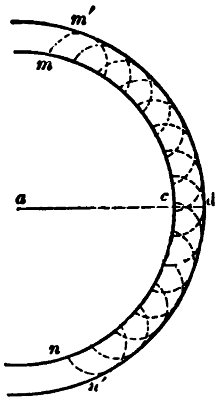


Fig. 113

194. Huyghens' Principle. — Let a be the center of disturbance, and let it be surrounded by a surface or wave front (only part of which, *mcn* is shown in Fig. 113). Then it is clear that the motion in the medium outside this surface must be fully determined by the motion existing for the moment in the wave front *mcn*. The single wave as it travels outward disturbs all the elements of the medium through which it passes. The disturbance of any element may then be considered as one cause of the subsequent disturbance of all the other elements. Hence Huyghens' principle that every point of the wave surface becomes a new center of disturbance, from

which waves are propagated outward in the same manner as from the original center, and the aggregate effect at any point outside the surface *mn* is the resultant of all the combined disturbances due to the secondary waves from these new centers. Every particle on the wave surface *mn* has the same oscillatory motion, except in point of amplitude, as the first particle disturbed; it therefore stands in the same relation to adjacent particles, and communicates motion to them in the same manner, or becomes itself a center of disturbance.

The principle of Huyghens is the principle of superposition. The disturbance at any point is due to the superposition of all the disturbances reaching it at the same instant from the various points of the medium through which the wave front passed an instant earlier.

Let the points of the surface *mn* be centers from which waves proceed for a short distance *cd*. Then with these centers and a radius *cd* describe circular waves. The number of such waves being indefinitely large, they ultimately coalesce to form the new surface $m'n'$, which is the envelope of all the small secondary waves. The effective part of each secondary wave Huyghens supposed confined to that portion which touches the envelope. In fact Stokes showed that the secondary waves mutually destroy one another, except at the surface enveloping them.

The energy of *mn* is thus passed on to $m'n'$, and in the same form or manner from $m'n'$ to $m''n''$, etc.

195. Reflection of a Plane Wave at a Plane Surface.—Let *CD* (Fig. 114) be the reflecting surface and *AB* a portion of the plane advancing wave front. To find the relation between the inclination of the incident wave front and that of the reflected wave front to the reflecting surface proceed as follows:

When the point *A* reaches the reflecting surface, it becomes by Huyghens' principle a new center of disturbance traveling back into the first medium. Then with *A* as a

center and with a radius AA' equal to BB' describe a circle. This circle serves to limit the distance traveled by the reflected disturbance while the disturbance from B is traveling to B'. In the same time the disturbance from b would have reached b' if there had been no obstruction. But in point of fact it travels to E and is there reflected. Hence with E as a center and with a radius Eb' draw another circle. In the same manner draw any convenient number of circles.

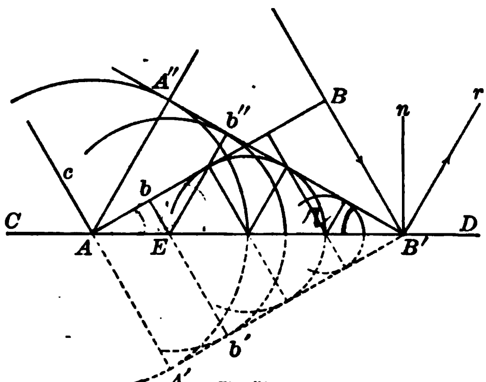


Fig. 114

Finally from B' draw a tangent to the first circle; it will touch all the other circles and will be the reflected wave front. Draw AA'' to the point of tangency on the first circle. AA'' is the path of the disturbance reflected from A and is therefore a ray. The triangle $AA''B'$ is symmetrical with $AA'B'$ and equal to it; it is therefore equal to the triangle ABB'. Hence the angles BAB' and $A''B'A$ are equal to each other. But the former is equal to the angle of incidence and the latter to the angle of reflection, for they are equal respectively to $BB'n$ and $nB'r$, and these are the angles made by the incident and reflected rays with the normal at the point of incidence. Hence, *the angle of incidence is equal to the angle of reflection.* The former is the angle between the incident wave front and the reflecting surface; the latter is the angle between the wave front after reflection and the reflecting surface.

196. Relation of Direct and Reflected Systems of Waves.—Let O (Fig. 115) be the origin of the incident spherical waves,

and let AB be the plane reflecting surface. Without reflection these waves would take positions at equal successive intervals of time indicated by the dotted lines; but because they are reflected, their positions by the Huyghens' principle are the full line curves symmetrically situated in front of the reflecting surface.

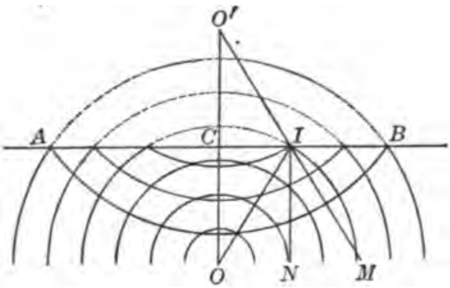


Fig. 115

Let OI be the path traveled by the disturbance from any point of the incident wave. Draw IM so that OI and IM make equal angles with the normal IN. Then IM is the path of the reflected disturbance: that is, it is normal to the reflected waves. Project IM backwards until it intersects at O' the normal through O to the reflecting surface. Then O' is the center of the reflecting waves. The triangles OIC and $O'IC$ are equal. Therefore OC and $O'C$ are equal, and the centers of the incident and reflected waves lie on a normal to the reflecting surface and are equidistant from it.

197. Stationary Waves.—When we are dealing with a train of waves, water waves for example, and not with a single wave, any point in the liquid may be affected at the same time by an incident wave and by one traveling in the opposite direction after reflection. The result of two systems of equal waves traversing the same medium in opposite directions is a system of *stationary waves*.

Let A (Fig. 116) be a wave moving to the right and B a wave of the same frequency and amplitude moving to the left. At points b, c, d, etc., midway between the points of zero displacement of the two waves, the displacement of the wave A is equal and of opposite sign to that of wave B.

Hence, *b*, *c*, *d*, etc., are points of zero displacement of the medium and of the resultant wave. Also bb' and bb'' will increase and decrease together equally because the two waves are equal and travel with the same speed in opposite directions. These points therefore remain at rest and are called *nodes*. Intermediate points vibrate from zero displacement

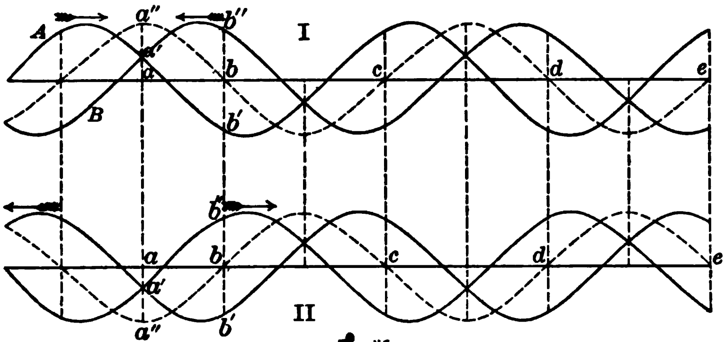


Fig. 116

to a maximum, first in one direction and then in the other. Points midway between the nodes and having the maximum amplitude of vibration are called *antinodes*.

In II each wave has moved forward half a wave length as compared with I. The nodes remain fixed, but the stationary wave has its displacements at all points of opposite sign to those of corresponding points in I. There is no progressive movement of the crests and troughs of the resultant wave, but only a periodic increase and decrease of amplitude. Such waves are therefore properly called stationary waves.

CHAPTER VI

PRODUCTION AND TRANSMISSION OF SOUND

I. NATURE AND MOTION OF SOUND

198. Sound and Hearing.—The word *sound* is commonly used in two distinct senses : first, to designate the sensation produced when a disturbance is conveyed to the brain by the auditory nerves ; and, second, the external cause of that sensation. These are the subjective and objective aspects respectively of the phenomena of sound. The external stimulus stands first in the series of energy changes leading to a sensation, but it is not *like* the sensation. All the external phenomena of sound may be present without the hearing ear. Sound should therefore be distinguished from hearing in the study of sound.

Sound may be defined as that form of vibratory motion excited in an elastic body which affects the auditory nerves and produces the sensation of hearing.

199. The Source of Sound a Vibrating Body. — The source from which sound proceeds is always a body in a state of vibration. Sound and vibratory movement are so related that one is strong when the other is strong, and they diminish and cease together.

If a mounted tuning fork is sounded and a light ball of pith or ivory, suspended by a thread, is brought in contact with one of the prongs at the back, it will be briskly thrown away by the energetic vibration of the fork.

Partly fill a goblet with water and produce a musical note by drawing a bow across its edge. The tremors of the glass will be communicated to the water and will throw its surface into violent agitation in four sectors, with intermediate areas of relative repose. This agitation disappears as the sound ceases.

A stout glass tube, four or five feet long, may be made to emit a musical sound by grasping it by the middle and briskly rubbing one end with a cloth moistened with water. So energetic are the longitudinal vibrations excited that it is not difficult to break the tube near the hand, on the side opposite to the end rubbed, into many narrow rings.

Any regular succession of taps produces a musical sound. The taps must be regular or periodic to make the sound musical; otherwise it is only noise. The vibrating body producing sound may be solid, liquid, or gaseous. Only the first and last are employed in musical instruments, the first comprising all instruments with strings, reeds, or bars, and the last all wind instruments.

200. Transmission of Sound to the Ear. — Sound requires for its transmission to the ear an uninterrupted series of elastic bodies. They may be solid, liquid, or gaseous. If the vibrating source of sound be isolated so that there is an interruption in the elastic medium, the vibrations do not reach the ear and no sound is perceived. Hence the well-known experiment of suspending a bell by a thin string within the receiver of an air pump. The sound is greatly enfeebled by exhausting the air; and if hydrogen be then admitted and the exhaustion repeated, the sound will cease altogether, though the hammer may still strike the bell. The bell has then no elastic medium in contact with it, to which it can give up its energy of vibration.

The transmission of sound by a solid is illustrated by the acoustic or string telephone. It consists of a taut string connecting the centers of thin elastic membranes stretched over the bottom of two small conical boxes. When one speaks or even whispers into one of these boxes, a listener at the other can hear distinctly, even at a distance of several hundred feet. The membrane vibrating transversely sets up longitudinal vibrations in the stretched string. These are transmitted to the membrane of the receiving instrument, and its motions reproduce the sounds actuating the transmitter.

Sound waves consist of a series of condensations and rarefactions, succeeding each other at regular intervals. Each particle of air vibrates in a short path in the direction of the sound transmission. Its vibrations are *longitudinal* as distinguished from the *transverse* vibrations in water waves.

201. Motion of the Particles and of the Wave. — The motion of the particles of the medium conveying sound is quite distinct from the motion of the sound wave itself. This distinction holds for all undulations transmitted through a medium of motion. A sound wave is composed of a condensation followed by a rarefaction. In the former the particles have a forward motion in the direction in which the sound is traveling; in the latter they have a backward motion, while at the same time both condensation and rarefaction travel steadily forward with a speed nearly independent of that of the particles composing the wave.

The independence of the two motions is aptly illustrated by a field of grain across which waves excited by the wind are coursing. Each stalk of grain is securely anchored to the ground while the wave sweeps onward. The heads of grain in front of the crest are found to be rising, while all those behind the crest and extending to the bottom of the trough are falling. They all sweep forward and backward, not *simultaneously*, but *in succession*, while the wave itself travels continuously forward.

A series of particles along the line in which a wave is traveling are in successively different phases of motion; and the distance from one particle to the next one *in the same phase* is a wave length. While any element of the medium merely oscillates about its position of stable equilibrium, there is a continuous handing on or flow of energy from point to point.

202. Characteristics of Musical Sounds. — Musical sounds differ from one another in three important particulars:

1. *Pitch.* The pitch of a note depends on the frequency or number of vibrations per second. An acute note has a greater frequency than a grave one. Also, since with a given speed of transmission the wave length is inversely as the frequency, a note may be designated by its wave length as well as by its frequency.

2. *Loudness.* The loudness of a sound depends on the energy of the vibrations transmitted to the ear. It involves also the pitch of the note. The energy of vibration is pro-

portional to the square of the amplitude; but as it is obviously impracticable to express a sensation with mathematical precision, it will suffice here to say that the loudness of a note increases with the amplitude of vibration.

3. *Quality.* Two notes of the same pitch and loudness, such as those of a piano and a violin, are yet clearly distinguishable by the ear. This distinction is expressed by the term *quality* or *timbre.* Helmholtz demonstrated that the quality of a note is determined by the presence of tones of higher pitch, whose frequencies are simple multiples of that of the fundamental or lowest tone.

Pitch depends on the *length* of the sound wave, loudness on its *amplitude*, and quality on the *form* of the complex wave.

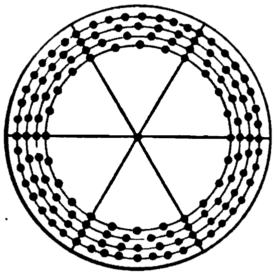


Fig. 117

203. The Siren. — An instrument in which a note is produced by the escape of air at regular intervals through holes in a rotating disk is called a *siren.* Thus, if the perforated disk of Figure 117 be mounted on the shaft of a small electric motor and rotated at a constant speed, a musical note may be produced by blowing a current of air through either of the concentric series of equidistant holes. The pitch will be different for each circle of holes, and will rise and fall with increase and decrease of speed.

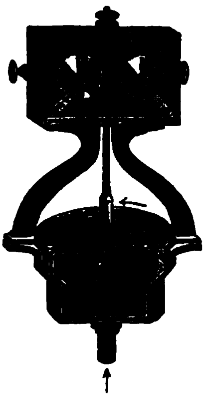


Fig. 118

The siren of Caignard de la Tour is shown in Figure 118. The rotating portion is mounted on a wind chest, which has a top perforated with a circle of equidistant holes. The revolving plate *D* contains the same number of holes as the lid *C*, and a puff of air issues simulta-

neously from all the holes every time those in the disk coincide with those in the lid. For each rotation of the disk there are thus as many puffs as there are holes in the circle. The rotations of the disk are counted by means of a gear train similar to that of a gas or electric meter.

The rotation of the disk and spindle of the siren of Figure 118 is produced by air pressure. The holes in both lid and disk are drilled obliquely and slope in opposite directions, as shown at *E*. Hence, a certain pressure is exerted against one side of the holes sufficient to set the disk in rotation.

The frequency of a given note may be determined by keeping the speed as nearly constant as possible, and throwing the counting train in gear with the spindle for an observed length of time. Then the product of the number of revolutions and the number of holes in the disk divided by the time in seconds gives the frequency of the note produced by the siren.

II. VELOCITY OF SOUND

204. Velocity of Sound in Air. — Common experience shows that the velocity of sound in air is a very moderate quantity. Thus, when a man is seen at a distance striking a blow with a hammer, the observer *sees* the blow struck an appreciable interval of time before he *hears* the sound. The time taken by light to travel over the intervening distance is entirely inappreciable in comparison with that of sound.

Again, one may hear a sound produced at a distance both through a telephone and through the air, the latter report arriving for a distance of 1100 feet about a second later than the former.

When one utters a loud sound a few hundred feet in front of the plane surface of a large building, the sound returns as a reflected wave or echo after an interval of time required for the sound to travel to the reflecting surface and back again. This interval is the same as the sound wave would take to travel to the observer from the center of the reflected

system of waves as far behind the reflecting surface as the observer is in front (§ 196).

Either of these three methods of observation may be made the basis of an experimental determination of the velocity of sound in free air. The problem, however, is beset with many difficulties. In the open air the disturbing effects of the wind and of local differences in temperature are beyond the control of the observer. The velocity of sound is affected also by the amount of moisture in the atmosphere. To these disturbances must be added the delay in sense perceptions after the arrival of the exciting cause of them. This would introduce no error if the delay or reaction periods were the same for sight and hearing; but they are not the same, and neither is the same for different persons.

The effect of the wind has been eliminated by firing a cannon from the two stations in succession, and taking the mean of the intervals between seeing the flash and hearing the report.

Beginning with the determinations made by the French Academy of Sciences, the following are the most trustworthy results, all reduced to 0° C.:

1.	Academy of Sciences, 1738	332.00 m./sec.
2.	Bureau des Longitudes, 1822	331.00 m./sec.
3.	Moll and Van Beck, 1823	332.25 m./sec.
4.	Bravais and Martins, 1844	332.37 m./sec.
5.	Stone, 1871	332.40 m./sec.

In Stone's experiments in South Africa two observers were stationed at unequal distances from a time gun, and each gave a signal with an electric key on hearing the report. This method eliminated the eye observations entirely, and confined them to the ear. The observers were stationed respectively 641 feet and 15,449 feet from the gun. The signals were recorded on a chronograph at the Observatory in Cape Town. Reciprocal firing was not resorted to, but allowance was made for the wind. The results were cor-

rected also for the difference in the personal equation of the observers, — the time required to perceive and record an event. The difference in the distances of the two observers from the time gun divided by the interval recorded on the chronograph gave the velocity of sound. The final value obtained by Stone for 0° C. was:

$$V_0 = 332.4 \text{ m./sec., or } V_0 = 1090.6 \text{ ft./sec.}$$

205. Regnault's Experiments. — In the years from 1862 to 1866 Regnault made an extensive series of observations on the propagation of a sound pulse through water pipes in Paris, in lengths up to 4900 m., and of diameters from 10.8 cm. to 1.1 m. When a gun was used as a source of sound, the instant of firing was recorded electrically by the breaking of a wire stretched across the muzzle; the interruption was recorded on the revolving drum of a chronograph. The receiver at a distant point was a thin rubber membrane stretched over the smaller end of a wide cone. The movement of this membrane or drum was made to close an electric circuit, and the arrival of the sound pulse was thus recorded on the cylinder of the chronograph.

Since the time interval was recorded automatically, it might at first appear that the error due to the personal equation was eliminated; but in fact such a membrane has as much a personal equation as an observer. Time is required by it to acquire sufficient energy from the wave to move the membrane, and there is thus an inevitable delay in responding to the sound. This delay was the greater the feebler the sound. Regnault sought to allow for it, but it is probable that his results are slightly under the true value.

Regnault found that the velocity of sound decreases with its intensity, tending toward a lower limit for feeble sounds. Also, that it increases with the diameter of the tube, tending toward an upper limit in very wide tubes. His final result for a feeble sound in a wide tube at 0° C. was:

$$V_0 = 330.6 \text{ m./sec., or } V_0 = 1084.6 \text{ ft./sec.}$$

206. The Krakatoa Eruption. — At a temperature only a little above 0° C. the velocity of sound is 1100 ft./sec. Compare this with the speed of a railway train:

88 ft./sec. = 5280 ft./min = 1 mile/min. = 60 miles/hour.

Add one quarter, and

110 ft./sec. = 6600 ft./min = 1.25 miles/min. = 75 miles/hour.

Sound at 1100 ft./sec. is then only ten times the speed of a very fast railway train, or 750 miles per hour. The air-line distance between New York and Chicago is about 750 miles. Hence a very loud sound would travel from one city to the other in an hour.

The Krakatoa eruption, August, 1883, furnished a very remarkable example of sound waves on a gigantic scale. A series of explosions culminated in a tremendous outburst, which blew a part of a mountain into the air. At a distance of 2000 miles these explosions sounded like the firing of heavy guns. The chief outburst produced a pulse or wave of such intensity that it left a noticeable trace on self-recording barometers as it traveled around the globe. A study of these records showed that the wave traveled with a speed of about 700 miles an hour. It culminated on the side of the earth opposite Krakatoa in 18 hours; it then spread out again and culminated several times afterward, its trace not being entirely lost until after 127 hours.

207. Theoretical Determination of the Velocity of Sound. — The velocity of a longtitudinal wave may be calculated in terms of the coefficient of volume elasticity k and the density d of the medium.* The result

$$V = \sqrt{\frac{k}{d}} \tag{47}$$

was first deduced by Newton. This formula applies directly to transmission through solids and liquids. For gases it may be modified by substituting the pressure of the medium

*Carhart's *University Physics*, Part I, p. 152.

for the coefficient of elasticity; for by Boyle's law, if P and S are the original pressure and volume and $P+p$ and $S-s$ the corresponding new values, then

$$\begin{aligned} PS &= (P+p)\,(S-s) \\ &= PS + pS - Ps - ps. \end{aligned}$$

Neglecting the product ps of two very small quantities, we have

$$Ps = pS, \text{ or } P = p\frac{S}{s} = k \text{ (§ 125).}$$

Hence, finally, substituting P for k,

$$V = \sqrt{\frac{P}{d}}. \tag{48}$$

$P = 1{,}013{,}250$ dynes per cm.^{2} and $d = 0.001293$. Therefore at 0° C. $V_0 = \sqrt{\frac{1013250}{0.001293}} = 27{,}993$ cm./sec., or 918.4 ft./sec. Direct observation gives 33,240 cm./sec. The discrepancy is about one sixth.

208. Laplace's Correction. — Newton attempted to explain the failure of his formula for the velocity of sound in air to give a result corresponding with the observed value; but his explanations were mere hypotheses, and it remained for Laplace to detect the error in 1816 and to point out the correction.

The proof that the coefficient of elasticity of the air is numerically equal to the pressure depends on an application of Boyle's law (§ 207); but Boyle's law is true only under the condition of a constant temperature. The corresponding coefficient is called the *isothermal* coefficient of elasticity of a gas. But there is another coefficient called the *adiabatic* coefficient of elasticity. This is the coefficient when no heat leaves or enters the gas during the compression and expansion

respectively. Now heat is generated in compressing a gas, and is absorbed when a gas expands. The compression and expansion of a gas are, therefore, not isothermal except when these changes take place very slowly so that equalization of temperature is possible. If they occur rapidly, there is no time for the temperature to come to equilibrium by conduction and radiation.

In sound waves the compressions and rarefactions follow each other in such quick succession that there is no time for the compressed portion of a wave to share its excess of heat with the cooler rarefied portion. The result of this rise of temperature in compression is that the pressure for any given decrease in volume must be higher than when the temperature is constant. So also the fall of temperature in rarefaction results in a greater decrease of pressure for a given increase in volume. The coefficient of elasticity is therefore greater than when these changes are isothermal. The relation between pressure and volume is not then the relation expressing Boyle's law

$$pv = \text{a constant } (\S\, 167),$$

but it is expressed by the formula of Poisson,

$$pv^{\gamma} = \text{a constant}.$$

The exponent γ is the ratio of the specific heat of a gas under constant pressure to its specific heat at constant volume (§ 390). The corresponding adiabatic coefficient of elasticity is γ times the isothermal coefficient, or γP.

For air $\gamma = 1.405$. Hence

$$V_0 = \sqrt{1.405 \, \frac{P}{d}} = 33{,}181 \text{ cm./sec.}$$

This value agrees very closely with the best experimental determinations and appears to justify Laplace's supposition that the changes of volume are adiabatic.

209. Correction for Temperature. — Changes of pressure, such as a change in the barometric pressure, unaccompanied by changes of temperature, do not affect the velocity of sound in air; for P and d then vary in the same ratio, and their quotient P/d remains unchanged. But this is not the case when a change of pressure is due to a change in temperature. If we assume that the relation between pressure and temperature is expressed by the equation

$$P = P_0(1 + \alpha t),$$

where α is the coefficient of expansion of a gas, $\frac{1}{273}$ or 0.00367, then the velocity at any temperature becomes

$$V = \sqrt{1.405 \frac{P_0\,(1 + \alpha t)}{d}},$$

or

$$V = V_0\,\sqrt{(1 + 0.00367\,t)}.$$

For small values of t this expression becomes

$$\begin{aligned} V &= V_0(1 + \tfrac{1}{2}\,\alpha t) \\ &= V_0\,(1 + 0.00183\,t). \end{aligned}$$

The increase in the velocity of sound for 1° C. is therefore $33{,}200 \times 0.00183 = 60.7$ cm./sec., or 23.9 in./sec.

210. Velocity of Sound in Water. — The compressibility of liquids is very slight and their isothermal and adiabatic coefficients of elasticity are substantially the same. Therefore the general formula for the velocity of sound $V = \sqrt{\frac{k}{d}}$ is directly applicable to them. The compressibility of water at 4° C., that is, the decrease in volume of a unit volume due to an increase of pressure of one atmosphere, is 0.0000499. Hence

$$k = \frac{\mathit{stress}}{\mathit{strain}} = \frac{1{,}013{,}250}{0.0000499}.$$

The density of water at 4° C. is unity. Therefore the velocity of sound in water at 4° C is

$$V = \sqrt{\frac{1{,}013{,}250}{0.0000499}} = 142{,}500 \text{ cm./sec., or } 1425 \text{ m./sec.}$$

In 1827 Colladon and Sturm measured with much care the velocity of sound in the water of Lake Geneva between boats anchored at a distance apart of 13,487 m. The mean time required for the transmission over this distance of the sound of a bell struck under water was 9.4 seconds. This gives for the observed velocity of sound in water 1435 m. at 8.1° C.

A system of transmitting signals through water by means of submerged bells is in use by vessels at sea and for offshore stations. Special telephone receivers have been devised to operate under water and to respond to the sound signals. Indeed, the vessel itself acts as a sounding board and as a very good receiver.

211. Reflection of Sound. — Whenever the medium transmitting sound changes suddenly in density, a part of the energy is transmitted and a part reflected. One system of waves then gives rise to two systems, and the intensity of the sound in either system is less than before reflection. The intensity of the reflected system is the greater the greater the difference in the densities of the two media. A dry sail reflects a part of sound and transmits a part; but when wet it becomes a better reflector and almost impervious to sound.

When an obstacle is interposed between the source of sound and the ear, if the dimensions of the obstacle are not much greater than the wave length of the sound, the waves close in around the edges of the obstacle and the intensity behind it is but little weakened. But for waves shorter than the dimensions of the obstacle, it casts a sound shadow, or the sound behind it is relatively faint. The dependence of the sound shadow on the dimensions of the obstacle as compared with the wave length may be strikingly shown in the following manner: Place the ear in line with a watch

and a clock, the watch much nearer the ear; then interpose a quarto book or magazine, or even a folded newspaper, between the ear and the watch. The ticking of the watch will become almost or quite inaudible, while that of the clock will be scarcely changed. The wave lengths of the ticking of the clock are perhaps ten times as long as those of the watch.

Reflection of sound may take place from the surface of a rarer medium as well as from that of a denser. Thus, it is easy to demonstrate that sound is reflected from a large flat gas flame, and that the angle of incidence is equal to the angle of reflection. In this case the phase of the reflected wave is changed by half a wave length, a condensation being reflected as a rarefaction and conversely.

When the reflection is from the denser surface, the motion of the air particles is reversed, and a condensation is reflected as a condensation; but when the reflection is from the surface of the rarer medium, the motion of the air particles is not reversed, each particle moving forward beyond its normal position of equilibrium, and a condensation is reflected as a rarefaction. Mechanically the impact of two unequal elastic balls illustrates the same phenomena. When the smaller ball strikes the larger one, its motion is reversed; but when the larger ball strikes the smaller, it moves forward without reversal, but with diminished velocity.

212. Echoes.—An echo is a sound reflected normally. A clear echo requires reflection from a vertical surface, the dimensions of which are large compared to the wave length of the sound. A cliff, a wooded hill, or the broad side of a building may serve as the reflecting surface. It must be smooth in the sense that its inequalities are small compared to the wave length of the incident sound. The same condition applied to the reflection of light requires highly polished surfaces for regular reflection as distinguished from diffusion.

Parallel reflecting surfaces at suitable distances produce multiple echoes, as parallel mirrors produce multiple images. For short distances from the reflecting surface, the direct and reflected sounds are confused, as in the case of rooms with bad acoustic properties. A circular room, like the Baptistry at Pisa, may prolong a sound for ten seconds or more by successive reflections. The effect at Pisa is made more conspicuous by the good reflecting surface of polished marble.

Echoes sometimes present peculiarities which may be referred to the character of the reflecting surface. The reflected sound, such as a shout, may be an octave higher than the incident sound. The octave is present in the shout itself, but is masked by the louder fundamental tone. The inequalities of the reflecting surface are such that the waves constituting the fundamental tone are diffused, while the shorter waves of the octave are regularly reflected with much smaller diminution of intensity.

Aërial echoes also are often observed. The reverberations of distant thunder are doubtless due in part to successive reflection from clouds. Air almost perfectly transparent to light may have great acoustic opacity. When for any reason the atmosphere becomes unstable, vertical currents are formed and vertical sections or banks of different densities. The sound transmitted by one bank is in part reflected by the next, the successive reflections giving rise to a curious prolonging of a short sound. Thus, the sound of a gun or a whistle is then heard apparently rolling away to a great distance with decreasing loudness.

213. Deflection of Sound Waves by the Wind. — Sound is heard better with the wind than against it. Sound waves traveling in the same direction as the wind are deflected downward, while those going against the wind are deflected upward. Imagine plane vertical waves traveling horizontally. The upper layers of the air move faster than those next to the ground; and if the wind is blowing in the same direction as

the sound is traveling, the velocity of the wind must be added to that of the sound. Hence the upper parts of the sound waves travel faster than the lower, the wave fronts are tilted forward, and there is a condensation along the surface of the ground, with increased loudness.

On the other hand, the velocity of sound against the wind is diminished by the wind, the upper parts of the waves are retarded more than the lower, the wave fronts are tilted backward, and the sound is deflected into the upper air with a consequent weakening near the ground.

A similar deflection of sound waves occurs when there is a vertical temperature gradient. The velocity of sound increases about 0.6 m. per degree C. If, therefore, there is a fall of temperature with increase in elevation, sound travels faster near the ground than higher up, the wave fronts are tilted backward and the sound may be deflected entirely over the listener. Such is the case on a hot summer day.

If the temperature gradient is reversed, the wave fronts are tilted forward, and the sound is condensed along the ground. When the air is still after sunset, it cools near the ground more rapidly than above, and sounds are then heard at great distances. Similarly on frosty mornings when the air is still, it is cooled by contact with the ground, and sounds carry remarkably well.

214. The Doppler Effect. — If the source of sound is approaching the ear with a velocity v, then more waves reach the ear in a second than if the source were stationary, and the pitch of the note is correspondingly higher.

The apparent wave length of the sound in the air is reduced by the distance traveled by the source during the period T of the note. This distance is vT or v/n. Then the apparent wave length becomes $V/n - v/n = 1/n\,(V - v)$, and the apparent frequency

$$V \div \frac{V - v}{n} = n\left(\frac{V}{V - v}\right).$$

Similarly, if the source is moving away from the observer, the apparent frequency, that is, the frequency reaching the ear, is $n\left(\frac{V}{V+v}\right)$. This phenomenon is known as the "Doppler effect." It may be readily detected if one notes the pitch of the whistle or of the bell of a locomotive when it approaches as compared with the pitch when it recedes in passing.

A similar phenomenon is observable in light when the source is moving with respect to the observer with a velocity comparable with that of light (§ 315).

Problems

1. How long will be required for sound to travel a distance of 2 mi. in air at 20°?

2. On a day when the temperature was 25°, the interval between seeing a flash of lightning and hearing the thunder was 4 sec. How far away was the lightning?

3. A gun is fired and after 4 sec. the echo from a distant hill is heard. How far distant is the hill, the temperature being 20°?

4. A shell fired at a target, half a mile distant, was heard to strike it 4.5 sec. after leaving the rifle. What was the average velocity of the shell, the temperature of the air being 20°?

5. If the velocity of sound in air at 20° is 1120 ft. per second, what would its velocity be in carbon dioxide at the same temperature, assuming carbon dioxide to be 1.44 times as heavy as air?

6. If sounds travel in air at the rate of 340 m. per second, and instantaneously through a telephone line, what would be the distance between the source and the observer if the interval between hearing the sounds by the two methods of transmission was 4 sec.?

7. If the velocity of sound in air is 340 m. per second, and in iron 5130 m. per second, what would be the interval between hearing a sound through an iron bar 340 m. long and through the air?

8. A tuning fork gives a note of the same pitch as a siren, which rotates 270 times in 10 sec., the disk having 16 holes. What is the frequency of vibration of the fork?

9. A locomotive is running at the rate of $\frac{10}{11}$ of a mile a minute. The difference between the apparent frequencies of vibration of the note given by the whistle when approaching the observer and after passing him is 37.5 vibrations per second. If the velocity of sound is 1120 ft. per second, what is the true frequency of the whistle?

10. A stone is dropped into a well and is heard to strike the water after 4 seconds. What is the depth of the well if the velocity of the sound is 335 m. per second?

11. If at night signals from a station are given simultaneously by means of a flash of light and a stroke on a submerged bell, what is the distance from the station of a ship on which the interval between seeing the flash and hearing the sound of the bell, by a telephone responding to the sound signal through the water, is $1\frac{1}{2}$ seconds, if the velocity of sound in the water is 1430 m. per second?

12. An observer is stationed at a distance of 6 km. from a gun, which is fired at noon; if the wind is blowing at a speed of 80 km. an hour, and the observer is stationed directly to windward, at what time will he hear the report?

CHAPTER VII

PHYSICAL BASIS OF MUSIC

I. MUSICAL INTERVALS AND SCALES

215. Musical Intervals. — The pitch of a musical note is the pitch of its gravest component, or fundamental tone.

Pitch may be defined in two ways:

1. *Physically*, as the number of vibrations per second in the fundamental tone of the note.

2. *Musically*, by referring the note to its place in an arbitrary scale of musical pitch.

When two notes are sounded together, the ear recognizes a relationship between them, known as *a musical interval.* Its measure is the *ratio of the frequencies* of the two notes. Names have been given to many such intervals employed in music. When the ratio is 1, the interval is called *unison;* 2, an *octave;* $\frac{3}{2}$, a *fifth;* $\frac{4}{3}$, a *fourth;* $\frac{5}{4}$, a *major third;* $\frac{6}{5}$, a *minor third;* $\frac{25}{24}$, a *chromatic semitone.*

Any three notes whose frequencies are as 4 : 5 : 6 form a *major triad*, and together with the octave of the lowest note, a *major chord.* Three notes whose frequencies are as 10 : 12 : 15 form a *minor triad*, and together with the octave of the lowest note, a *minor chord.*

A single tracing point may be set in motion by two or more systems of sound waves at the same time. If the surface on which the curve is to be inscribed is moved at right angles to the motion of the tracing point, the resultant curve will be due to the composition of the several motions in the same plane. The upper curve in Figure 119 is the result

of combining in this way three simple harmonic motions composing a major triad. The lower curve in the same way shows the composite

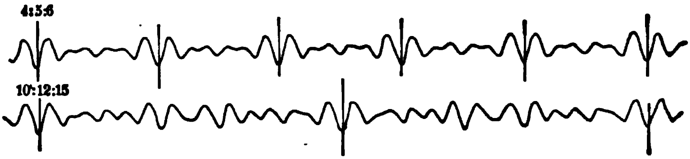


Fig. 119

motion resulting from a minor triad. It will be noticed that a complex wave recurs in both cases, but less frequently in the second than in the first.

216. The Diatonic Scale. — Continuity of tone from one note to the next, such as occurs in the whistling of the wind, or in the moaning of animals in pain, has always been avoided in musical compositions. The change of pitch in any melody takes place by steps or intervals, and not by continuous transition. A *musical scale* is the succession of notes by which musical composition ascends from one note, called the *keynote*, to its octave. The last note in one scale is regarded as the keynote of another series of eight notes with the same succession of intervals. In this way the series is extended until the limit of pitch established in music is reached.

The common succession of eight notes in the major mode, called the *diatonic scale*, is one of many used at different times, and finally adopted by European peoples about three hundred and fifty years ago on account of its fitness to express the style of music cultivated by them.

In the notation employed by Helmholtz the octave below bass c is written

$$C \quad D \quad E \quad F \quad G \quad A \quad B \quad c.$$

The octave above c is

$$c \quad d \quad e \quad f \quad g \quad a \quad b \quad c'.$$

The octave above c' (middle C) has one accent, and the next higher two accents.

In each octave there are three major triads for the key of C:

$$\left.\begin{array}{l} c' : e' : g' \\ g' : b' : d'' \\ f' : a' : c'' \end{array}\right\} :: 4 : 5 : 6.$$

It is the universal custom in Physics to take the frequency of c' as 256, because this number is a power of 2. This frequency is practically that of the middle C of the piano. If, then, c' is due to 256, or m, vibrations per second, the frequencies of the other notes of the diatonic scale may be found by simple proportion; they are as follows:

256	288	320	$341\frac{1}{3}$	384	$426\frac{2}{3}$	480	512
c'	d'	e'	f'	g'	a'	b'	c''
Do	Re	Mi	Fa	Sol	La	Si	Do
m	$\frac{9}{8}m$	$\frac{5}{4}m$	$\frac{4}{3}m$	$\frac{3}{2}m$	$\frac{5}{3}m$	$\frac{15}{8}m$	$2m$

If the fractions representing the vibration frequencies are reduced to a common denominator, the numerators may be taken to represent the relative frequencies of the eight notes of the scale. They are

24 27 30 32 36 40 45 48

The intervals from each note to the next higher are:

Major Tone	Minor Tone	Half Tone	Major Tone	Minor Tone	Major Tone	Half Tone
$\frac{9}{8}$	$\frac{10}{9}$	$\frac{16}{15}$	$\frac{9}{8}$	$\frac{10}{9}$	$\frac{9}{8}$	$\frac{16}{15}$

The interval between a major tone and a minor tone is $\frac{81}{80}$; it is the smallest interval recognized in music and is called a *comma*.

217. Transposition. — In addition to the eight notes of the diatonic scale, use is made of additional notes derived from the eight by raising or lowering the pitch of each by a chromatic semitone, $\frac{25}{24}$. If the pitch is raised, the note is

said to be ***sharp;*** if lowered, ***flat.*** Thus *G*♯ is read *G* sharp, and *G*♭ is read *G* flat.

Sharps and flats are introduced by changing scales. It is necessary to have a choice of scales with keynotes of different pitch to suit different voices. Also, to avoid monotony in the same composition, it is desirable to be able to change from one scale to another, whose keynote bears a simple relation in frequency to the keynote of the first scale. Hence arises a series of scales whose keynotes are the successive notes of the first one. Many notes are common to all these scales, and when negligible intervals are eliminated, only twenty notes remain to each octave. The process of changing from one keynote to another is called *transposition.*

If, for example, the keynote is transferred from *C* to *G*, then the eight notes of the new scale must follow *G* with the same succession of intervals as in the key of *C* (§ 216). The only notes in the key of *G* differing from those in the key of *C* are *F* and *A*. A numerical comparison of the two scales shows the precise difference.

Key of *C*

c'	d'	e'	f'	g'	a'	b'	c''	d''	e''	f''	g''
256	288	320	$341\frac{1}{3}$	384	$426\frac{2}{3}$	480	512	576	640	$682\frac{2}{3}$	768

Key of *G*

256	288	320	360	384	432	480	512	576	640	720	768

The interval between the *a*'s of the two scales is only $\frac{81}{80}$, and may be neglected. The interval between the *f*'s is much larger and in the key of *G* the *f* becomes *f*♯. Instead of writing the sharp with this note every time it occurs, it is usually placed at the beginning of the staff and forms what is called the "signature" of the key. The signature of the key of *G* is therefore *F* sharp.

218. Tempered Scales.—Every transposition from one key to another increases the number of notes in an octave, and the number required for just intonation in all keys is in excess of what is mechanically practicable in instruments with fixed keys, such as the piano and organ. It is therefore necessary to adopt some system of accommodation by which the advantages of just intonation are sacrificed in order to reduce the number of notes. Such an adjustment or compromise is called *temperament*.

By sacrificing the advantage of notes with frequencies in simple ratios to the keynote, in order to avoid the practical difficulties of musical execution, it has been found possible to make a sufficient variety of scales out of twelve notes to the octave. Since the interval from E to F and from B to C in the diatonic scale is already a semitone, no other note is interpolated at those points. The extra notes occur where there are whole tones, or in groups of twos and threes, represented by the black keys on the piano, making thirteen notes to the octave and twelve intervals.

The system of temperament most commonly applied to the piano and organ is the system of *equal temperament* introduced by Bach. This system ignores the difference between major and minor tones, and makes all the intervals from note to note the same. This ratio applied twelve times equals an octave, or 2. The half-tone interval in the system of equal temperament is therefore equal to $\sqrt[12]{2} = 1.05946$. The results differ widely from pure intonation. The only accurately tuned interval in such a scale is the octave, all the others being more or less modified.

The numbers following show the differences between the diatonic and equally tempered scales:

	c'	d'	e'	f'	g'	a'	b'	c''
Diatonic	256	288	320	341.3	384	426.7	480	512
Tempered	256	287.3	322.5	341.7	383.6	430.5	483.3	512

The frequencies of none of the corresponding notes differ as much as one per cent. In a melody where only single notes are sounded in succession, the difference between the two scales is hardly noticeable; but when several notes are sounded in a chord, the contrast is more marked.

219. Musical Pitch. — The intervals between the notes of the scale are quite independent of the absolute frequencies, and depend on the ratios of these frequencies only. The actual pitch employed in music has varied widely in the last two centuries. In the time of Handel middle *A* (a') had a frequency of 424 vibrations per second, while for the organ in England in the middle of the eighteenth century a' was as low as 388 vibrations. In Paris in 1700 a' was 405, but in 1857 it had gradually risen to 448. As late as 1857 a maker of pianos in New York tuned his pianos by an a' fork giving 451.7 vibrations per second. Modern concert pitch has risen as high as 460 for a'.

There are several possible reasons for this progressive rise of pitch; the most obvious one is that instrument makers have intentionally raised concert pitch for the purpose of increasing the brilliancy of orchestral music. At the same time vocalists have been made to suffer.

The frequency 435 for a' was chosen by the Paris Academy of Sciences, and at present probably this pitch receives the widest recognition. In 1834 the German Society at Stuttgart recommended $a' = 440$ as the standard pitch. This makes $c' = 264$, or 11 times 24. The notes of the diatonic scale for the key of *C* corresponding to this standard may accordingly be found by multiplying the numbers 24, 27, 30, 32, 36, 40, 45, 48, by 11. The corresponding frequencies for the diatonic and equally tempered scales are then the following, the a's being the same in both:

	c'	d'	e'	f'	g'	a'	b'	c''
Diatonic	264	297	330	352	396	440	495	528
Tempered	261.6	293.7	329.6	349.2	392	440	493.9	523.3

220. Limits of Pitch. — In the modern piano of seven octaves, the bass *A* has a frequency of about 27.5, the highest *A*, 3480. Allowing for slight variations from the standard in tuning, the range of frequency for the piano is from 27 to 3500 vibrations per second.

The gravest note of the organ is the C of 16 vibrations per second, given by the 33-foot open pipe. Its wave length in air at normal temperature is $\frac{344}{16} = 21.5$ m., or 70.5 ft. The highest note is the same as the highest A of the piano, the third octave above a'.

The cultivated voice of a singer has a range of about two octaves. The voice of women has about twice the frequency of that of men. The lowest note of the human voice, not including certain exceptional cases, is C of 65 vibrations. The entire range is included between this note and $c''' = 1044$, the two higher octaves belonging to women.

The limits of hearing far exceed those of music. The range of audible sounds is about eleven octaves, or from $C_2 = 16$ to $c^{\text{VIII}} = 32{,}768$, though many persons of good hearing perceive nothing above $c^{\text{VII}} = 16{,}384$. By means of a Galton's whistle of adjustable length, a series of short waves may be produced, gradually passing beyond the upper limit of hearing.

221. Resultant Tones.—When two loud notes, differing less than an octave, are sounded together, they give rise to a third tone, called a *resultant tone*, whose frequency is the difference of the vibration rates of the two. They were discovered by Sorge in 1740, and independently by Tartini in 1754; they are therefore often called Tartini's tones.

Two notes c'' and g'', whose frequencies are 512 and 768, together give a distinct tone c' with a frequency of $768 - 512 = 256$. So also c'' and e'', with frequencies 512 and 640, give as a resultant tone $640 - 512 = 128$, or c.

Helmholtz called these tones *difference tones* to distinguish them from the *summation tones*, discovered by himself. The frequency of the latter is equal to the sum of the frequencies of the two notes producing them.

Summation tones are fortunately much feebler than difference tones, since they are mostly inharmonic.

II. TRANSVERSE VIBRATION OF STRINGS

222. The Sonometer.— In sound a *string* signifies a cord or wire stretched between two fixed points with such a tension that when the string is slightly displaced and released, it vibrates to and fro, and gives a musical note. For small displacements the increased force due to the flexure of the wire or cord is usually negligible in comparison with the restoring force due to the tension.

The *sonometer* consists of a thin wooden box, across which is stretched a violin string or more commonly a thin piano wire (Fig. 120). The wire is fixed at one end and passes

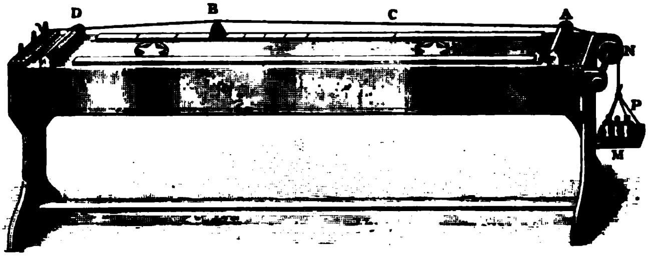


Fig 120

over two bridges, A and B, near the ends, and finally over the pulley N. It is desirable to have a comparison wire which can be tuned by means of a wrest pin, or by another pulley with weights. By means of a movable bridge B, sliding along a scale, the length of the wire may be shortened at pleasure.

223. Laws of Strings.— The sonometer may be used to verify the following laws of strings:

1. *The vibrations are isochronous for a given string at a given tension.* If the string is plucked or bowed, the pitch of the note remains the same while its intensity dies away as the amplitude decreases. The pitch is therefore independent of the amplitude.

2. *The frequency of vibration is inversely as the length for a given tension.* Tuning the two wires to unison, shorten

one of them by moving the bridge B to $\frac{8}{9}, \frac{4}{5}, \frac{3}{4}, \frac{2}{3}$, etc. The successive intervals between the notes given by the two wires will be $\frac{9}{8}, \frac{5}{4}, \frac{4}{3}, \frac{3}{2}$, etc. Hence the notes given by the wire of variable length are those of the diatonic scale, and the frequencies are inversely as the lengths.

3. *When the length is constant, the frequency varies as the square root of the tension.* Starting with a given tension and the strings in unison, increase the stretching load on AD four times; it will then give the octave of the other. Increase it nine times, and AD will give the twelfth above the other with three times the frequency. These statements may be verified by dividing the comparison wire by a bridge into halves and thirds, so as to put it in unison with the other wire of variable tension.

4. *The length and tension being constant, the frequency varies inversely as the square root of the mass per unit length.* Stretch with equal tensions two strings differing in diameter and material, that is, in mass per unit length. Bring them to unison with the movable bridge. The ratio of the lengths will be the inverse of that of the square roots of the masses per unit length. If the strings are of the same material, their masses per unit length will be as the squares of their diameters.

The last three laws are all expressed by the following equation, in which n is the frequency of vibration, l the length of the string, t the tension, and m the mass per unit length:

$$n = \frac{1}{2l}\sqrt{\frac{t}{m}}. \tag{49}$$

The figure 2 in the denominator comes from the fact that for one vibration the disturbance traverses the length of the string twice, that is, the wave length in the string is $2l$.

224. Fundamental and Harmonics of a String. — When a string is plucked or bowed at its middle point, it vibrates as

a whole and gives its lowest or *fundamental tone.* But if a string is lightly touched or damped at its middle point, and either half is bowed, it then vibrates in two segments and yields the octave above the fundamental. If it is damped at a point one third of its length from one end, and the shorter portion is bowed, it will give the twelfth above the fundamental with three times its frequency. Damping in succession at points $\frac{1}{4}$, $\frac{1}{5}$, $\frac{1}{6}$, etc., from one end, and bowing every time the shorter length, notes of 4, 5, 6, etc., times the frequency of the fundamental will be obtained. The series of notes obtained in this way are called *harmonics of the fundamental.*

A simple addition to the above experiment serves to show that when a string is damped and bowed as described, it divides into vibrating segments, each equal in length to the one bowed, and separated by points

Fig. 121

relatively at rest. Cutting V-shaped slips of paper and placing them on the wire as *riders* (Fig. 121), it will be found that those at the middle of the segments are thrown off, while those at the points separating the segments remain seated. The two sets of riders are more easily distinguished if they are of different colors.

The intermediate points of least motion and the ends are called *nodes;* the vibration sections, *loops* or *segments;* the middle points of the segments, *antinodes.* The experiment illustrates *stationary waves.* The bowed segment sends waves along the wire which are reflected from the distant end, and the direct and reflected waves combine to produce stationary waves as described in § 197.

225. Melde's Experiment. — This experiment illustrates in a most beautiful manner the division of a string into equal

vibrating segments. A light silk cord is stretched horizontally between a small fixed pulley and one prong of a vertical tuning fork. In Figure 122 *A* the plane of the two branches of the fork contains the cord, the end of which is moved horizontally in the direction of its length when the fork vibrates. The cord relaxes, falls to its lowest position with the forward movement of the fork, again rises to the horizontal, and then to its highest position when the fork is again in its most forward position.

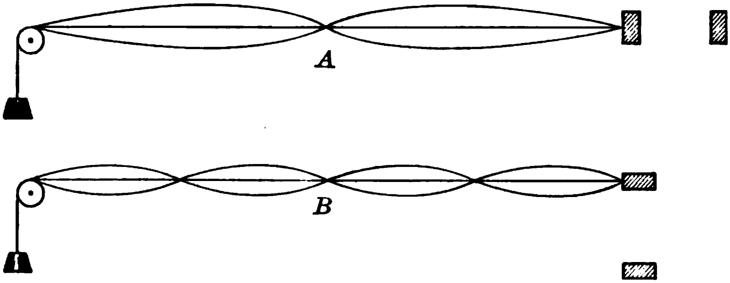


Fig. 122

The longitudinal movement of the point of attachment thus gives rise to a transverse vibration of the cord, with a period double that of the fork. If W is the weight necessary to make the cord vibrate as a whole, $W/4$ will make it vibrate in two segments, and $W/9$ in three segments. When the proper tension has been found, the cord spreads out in a pearl-white spindle, which appears to be fixed and stable.

If now the fork is turned on its axis (Fig. 122 *B*) so that it applies transverse impulses to the cord, the conditions then obviously require the fork and cord to vibrate in unison. Hence the cord vibrates in twice as many segments as in the first position; for if the frequency is doubled, the length of each segment must be halved with the same cord and stretching weight.

This experiment demonstrates the law of lengths and the

law of tensions. Moreover, from such experiments we conclude that,

A string may vibrate as a whole, or in any number of equal parts, the frequency being proportional to the number of parts.

226. Stroboscopic Observation of a Vibrating String. — The motion of a vibrating string is too rapid to permit it to be followed directly by the eye. A string may, however, be made to appear to execute its vibrations as slowly as one may desire by the "stroboscopic method" of illumination.

If the string is illuminated by intermittent flashes of light, the period between flashes being equal to the period of vibration of the string, it will be visible only in one of the positions through which it passes in a complete vibration, while in all other positions it is not illuminated and is invisible. Hence it appears to be at rest.

If the succession of flashes has a slightly different period from that of the string, they will find it in successive positions of its swing, and it will appear to execute a complete vibration in the period of time required for the intermittent illumination to gain or lose one flash as compared with the vibrations of the string. The more nearly the two periods coincide, the slower will be the apparent motion of the string.

Figure 123 illustrates one method of arranging the apparatus for the stroboscopic method of illumination. Sunlight reflected from the plane

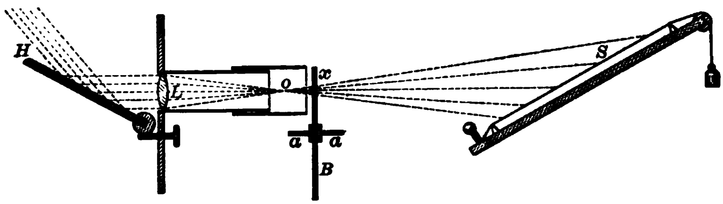


Fig. 123

mirror H passes through a converging lens L and is brought to a focus at O. Just beyond O is placed a perforated disk B, mounted on an axis of rotation aa. The disk should be perforated with from two to four holes on a circle about aa as a center and at equal distances apart.

Every time an opening comes to the position x, a flash of light passes through and illuminates the string S. The string is placed obliquely to the diverging beam of light, so that its whole length is illuminated without any great difference in intensity.

Instead of sunlight an electric lantern may be used; and the string is preferably horizontal and may be kept in vibration by a fork, as in

Melde's experiment. The experiment is the more interesting and instructive if the string vibrates in two or three segments.

The perforated disk may be mounted directly on the shaft of a small electric motor, and its rate of rotation may be adjusted by means of resistance in the electric circuit and by light controllable friction on the shaft. It is difficult to secure strictly uniform rotation of the disk, but its rotation may be kept constant long enough to show all the motions of the vibrating string.

The same method of illumination gives most beautiful results when applied to a spiral spring vibrating in segments. The vibrations are maintained by attaching the spiral spring at one end to the prong of a fork which is kept vibrating by an electromagnet. The spring is placed next to the condenser of the lantern and is projected through the openings in the disk in the usual way.

227. Coexistence of the Fundamental and Overtones.—If the wire of a sonometer be bowed vigorously and then be lightly damped at its middle point, the fundamental will disappear and the octave with twice the frequency will be heard instead. If again the wire be bowed as before and then be damped at one third of its length, it will cease to give the fundamental, and the twelfth with three times the frequency will probably be heard by an ear capable of recognizing pitch. When a string vibrates, it emits a series of tones, the lowest of which is the fundamental. All these are initially present, though more or less masked by the fundamental tone. The particular members present depend on the manner in which the vibrations are excited. A string may execute simultaneously all the modes of motion which it is capable of adopting singly. Hence the production of harmonics or overtones together with the fundamental.

The whole sound emitted by any source, such as a string, is termed a *note*. The trained musical ear assigns a pitch to this note, which is that of the lowest of the series; each irresolvable component of a note is called a tone, the lowest the *fundamental tone*, and the others *overtones* or *upper partial tones*. Overtones whose frequencies are exact multiples of the fundamental are called *harmonics*.

III. RESONANCE

228. Free Vibrations. — The vibrations which a body performs when it is disturbed and then left to itself, are called *free vibrations*. The period of such vibrations is the natural period of the system and is independent of the amplitude of vibration, provided only that the displacement is small. Thus, if a simple pendulum be drawn aside from its position of stable equilibrium and released, its vibrations are free. A ringing bell, a bowed tuning fork, or a plucked guitar string, left to itself executes free vibrations.

Free vibrations gradually decrease in amplitude, because in nature all motion is checked more or less promptly, the moving body giving up its energy to other bodies. The amplitude therefore gradually decreases to zero, while the period remains constant. Such vibrations are said to be *damped*.

229. Forced Vibrations. — A body is often compelled to surrender its own free period and to vibrate with more or less accuracy in a manner imposed upon it by an external force. When a periodic force is applied to an oscillatory system, and the system ultimately vibrates with a period the same as that of the force, the vibrations are said to be *forced*.

The two prongs of a tuning fork, with slightly different natural periods, mutually compel each other to adopt a common frequency. Huyghens discovered that two similar clocks, adjusted to slightly different rates, kept time together when they stood on the same table. The more rapid clock was delayed and the slower one quickened by mutual influence. These two cases are examples of mutual control, and the vibrations of both members of each pair are forced.

If one prong of a fork be strongly pulled intermittently by an external force so predominant that its period cannot be altered by any resistance offered by the fork, the fork will ultimately be forced to vibrate at a rate determined by the external force. Such is the case of a fork controlled

by an electromagnet operated by an intermittent current of a period slightly different from the natural period of the fork. The vibration at first is intermittent, because the vibrating system struggles to maintain its own period. The frequency of the intermittence is equal to the difference in frequency of the fork and the current. In time the intermittence dies away, and the fork adjusts itself in some way to the impulses of the current. The fork will maintain the imposed rate so long as it is compelled to do so, but it returns to its own normal frequency as soon as the current ceases. The nearer the rates of the forced and the free vibrations of a system agree, the wider is the amplitude. The amplitude reaches a maximum when the energy applied to the vibrating system exactly compensates for losses by friction and other causes.

230. Resonance.—The limiting case of forced vibrations, when the period of the external force is nearly or quite the same as the natural period of the oscillating system, is called *resonance.* The law of resonance is that vibrations will be taken up, and their energy absorbed, by any system capable of vibrating synchronously with them and exposed to their periodic impulses.

Resonance depends upon the cumulative effect of small disturbances applied to a vibratory system in such a way as to synchronize with its own motions. Thus, one string takes up the vibrations of another which has the same vibration rate. If two tuning forks, mounted on resonant boxes, are adjusted very exactly to unison, the vibrations of one will excite the other, even at a distance of several meters. If one is bowed in the vicinity of the other and then silenced, the second will be sounding loudly. Then after a short interval, damping the second, the first will be found to be again sounding, and so on.

Resonance may be mechanical without sound. A heavy weight suspended by a rope may be set swinging through a wide amplitude by tying to it a thread and pulling gently when the thread tends to slacken in the hand. Each effort then adds to the accumulated motion; the series of small impulses at the right intervals unite to produce a large amplitude of oscillation.

If two heavy pendulums, suspended side by side on knife edges on a slightly yielding stand, are carefully adjusted to swing in the same period, and one of them is set swinging, it will cause the other to swing. The one

initially at rest lags slightly in phase behind the other and absorbs its energy until the one first set in motion comes nearly to rest. The give-and-take process is then reversed. Many years ago a suspension bridge at Manchester in England was destroyed when a troop of cavalry was crossing it, the step of the troop keeping time with the natural swing of the bridge. Its swing reached an amplitude exceeding the limit of safety. It is now the custom to break step when bodies of either horse or foot cross a bridge.

Many striking cases of resonance are observed in the domain of sound; resonance is also of frequent occurrence in the phenomena of alternating electric currents, and resonance in that branch of Physics is quite as striking as in sound.

231. Air Resonators.—The resonant body is frequently a partly inclosed mass of air. Any hollow vessel, such as a tall vase, has a natural note of its own, which may be found by blowing softly across the edge at the top. A sea shell is a very good resonator, and "the sound of the sea" heard when such a shell is held close to the ear is a case of resonance. The mass of air in the shell has its own vibration rate, and it responds to any faint sound of the same frequency.

Fig. 124

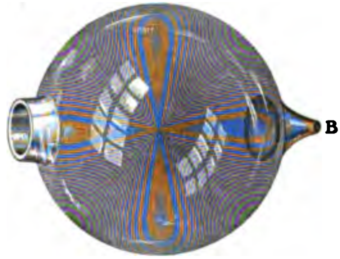
Hold an excited tuning fork over the mouth of a tall jar, the depth of which is greater than a quarter wave length of the note given by the fork (Fig. 124). Pour in water slowly; it will be found that at a certain level the note given by the fork is very greatly reënforced. Above or below that particular level, the intensity of the sound is not much affected by the presence of the column of air. Forks of different pitch require air columns of different length for marked reënforcement.

The box on which a tuning fork is mounted (Fig. 125) is a resonator tuned to the pitch of the fork and reënforcing its fundamental tone.

The spherical resonator (Fig. 126) was designed by Helmholtz for the purpose of picking out the overtones present in

Fig. 125

A



B

Fig. 126

a complex sound. The larger opening *A* is the mouth of the resonator; the smaller one *B* fits in the ear. When one of these resonators is held to the ear, it strongly reënforces a sound agreeing with it in pitch, but is silent to others.

IV. ORGAN PIPES

232. Air as a Source of Sound. — In many musical instruments, classed as "wind instruments," the sonorous body is a column of air in a pipe. Columns of air are set vibrating in two ways: by a vibrating tongue, as in reed instruments; or by a stream of air striking against the sharp edge of a lateral opening in the tube, as in the whistle, flute, and organ pipe.

Select several glass or metal tubes of different length and about 2 cm. in diameter; close one end with the palm of the hand and blow sharply across the edge at the other. Each pipe will give its own note, and the longer the pipe the graver the fundamental tone. If four tubes have lengths 32, 24, 20, 16 cm. respectively, the four notes will be c', e',

g', c'', forming a major chord. The lengths of the pipes are as 8, 6, 5, 4, and the frequencies of the notes as 4, 5, 6, 8; that is, *frequencies are inversely as the lengths of the pipes.*

233. Length of Pipe and Wave Length of the Fundamental Tone. — A comparison between the lengths of the pipes and the wave lengths of the notes given by them, as described in the last article, shows that a pipe closed at one end is approximately one fourth the wave length of its note in air. Middle C (c') of 264 vibrations per second has a wave length at a temperature of 20° C. of $\frac{34430}{264} = 130.4$ cm. The corresponding closed pipe has a length of $\frac{130.4}{4} = 32.6$ cm.

When experiments are made with a fork over pipes of different diameters, as in Figure 124, it is found that the length of a pipe resounding to a given fork is not quite constant, the length diminishing slightly as the diameter increases. When the length of the pipe is several times its diameter, the correction to be added to its measured length is 0.6 times the radius; that is, for a given frequency, length plus 0.6 radius is constant. Therefore the measured length of a pipe 2 cm. in diameter to give c' at 20° C. is $32.6 - 0.6 = 32$ cm.

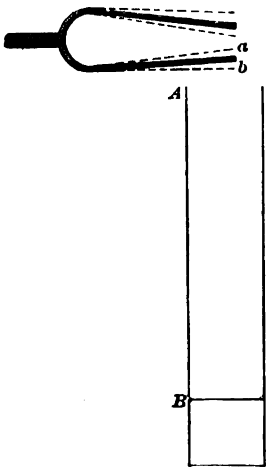


Fig. 127

When the prong at a (Fig. 127) moves to b, it makes half a vibration. It sends down the tube AB a condensation which is reflected at the surface B of the water and returns to the fork. If AB is $\frac{\lambda}{4}$, the distance down and back is $\frac{\lambda}{2}$, and the condensation returns to the

fork at the instant the prong reaches *b*. The prong next moves from *b* to *a* and sends a rarefaction into the tube following the condensation; the rarefaction returns to the open end of the pipe at the instant the fork arrives at *a* and completes its vibration. The initial relations are then reëstablished, and the air in the tall jar vibrates in synchronism with the fork. The disturbance has traversed the pipe four times while the fork has executed one complete vibration; hence a pipe closed at one end, called a stopped pipe, is one fourth the wave length of the fundamental tone emitted.

For the fundamental tone there is a node at the closed end of the pipe and an antinode at the open end. The greatest motion and least variation of pressure occur at the open end, and the least motion and greatest variation of pressure at the closed end. An excellent illustration of the motion of the air in a closed pipe is furnished by a spiral spring firmly fastened to a support at its lower end and attached to one prong of a fork at the other, the fork and spring being so adjusted that they vibrate in unison, with a node in the spring at the lower end only. The prong of the fork must vibrate in the direction of the length of the spring. The whole of the spring then opens out during one half of the vibration and closes in during the other half. The greatest variation of tension and the least motion occur at the fixed end of the spring, and the greatest motion and the least variation of tension at the end attached to the fork.

234. The Open Organ Pipe. — The open pipe has an antinode at each end, for there it is open to the air, and the air column at these points opens out most widely during vibration and undergoes the least change in pressure. But two antinodes never succeed each other without an intervening node. We should therefore expect to find a node at the middle of an open pipe for the fundamental tone.

The position of this node is readily found experimentally. Lower into an open pipe with one glass side a membrane

covered with fine sand (Fig. 128). When the pipe gives its fundamental tone, the sand is agitated least near the middle of the pipe and most at the ends.

Fig. 128

Since the open pipe has a node at the middle, it is equivalent to two stopped pipes with their closed ends together; its length is therefore half the wave length of the fundamental tone emitted by it.

To give the same note, the open pipe should be twice the length of the stopped pipe. If the correction 0.6 the radius be applied to each end of an open pipe and to the open end of the stopped pipe, which resounds to the same fork, the corrected lengths, within the range of experimental error, will be as two to one.

235. Overtones of Stopped Pipes. — The overtones of a stopped pipe are due to a division of the air column into segments, as represented in Figure 129. Each complete segment may be regarded as two pipes with their open ends turned toward each other, as represented in *B* and *C*. The presence of an additional node in a stopped pipe for the first overtone requires an additional antinode also. The half vibrating segment is then one third the whole length of the pipe as in *B*.

For the second overtone two nodes and two antinodes additional to those of the fundamental are required. The division of the air column is then as represented in *C*, with nodes at $\frac{1}{5}$ and $\frac{3}{5}$ the length of the pipe from the open end.

It is apparent that the overtones of the stopped pipe are the same as the fundamentals of pipes having lengths $\frac{1}{3}$, $\frac{1}{5}$, $\frac{1}{7}$, etc., that of the given pipe *These overtones are the odd harmonics with frequencies 3, 5, 7, etc., times the frequency of the fundamental tone.*

In adjacent half segments of an internodal space the mo-

tions are always in the same direction, or of the same sign. After a half period of the note produced, these motions are all again equal, but have changed sign. The motions of the air particles on the two sides of an inner node are always in opposite directions.

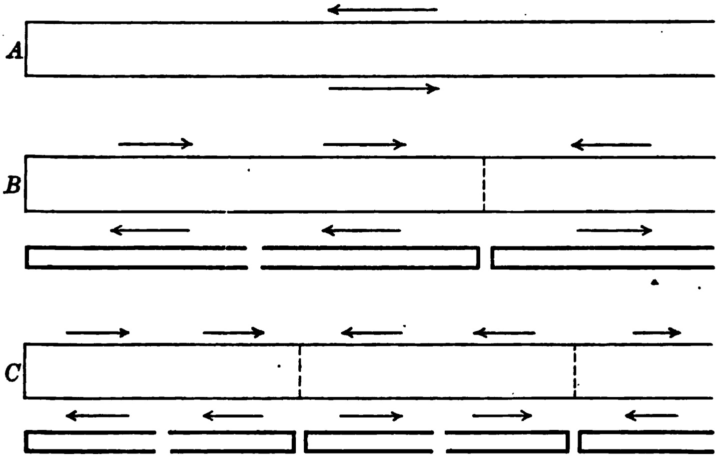


Fig. 129

236. Overtones of Open Pipes. — The overtones of open pipes are due to a division into segments of the vibrating column of air, as represented in Figure 130. For the fundamental the vibration in each half is like that in a pipe closed at one end. The motion of the air in the two halves of the pipe for a complete vibration is shown in *A*, the upper arrows for the first half vibration, and the lower ones for the second half. The air at the node in the middle of the pipe remains at rest, but is subject to variations of pressure due to alternate compressions and rarefactions.

The addition of another node and antinode gives four half segments as in *B*, instead of the two for the fundamental. The length of the segment for the first overtone is therefore half as great as for the fundamental, and the frequency is twice as great.

In C is shown the division of the air column for the second overtone. There are now six half segments of equal length instead of two, and each has a frequency of vibration three times that for the fundamental.

The overtones for an open pipe therefore form a complete harmonic series, with frequencies 2, 3, 4, etc., times the fundamental frequency.

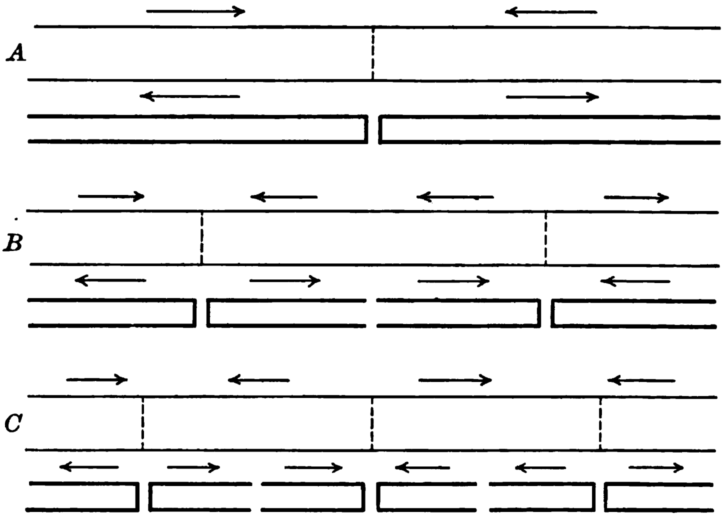


Fig. 130

237. Experimental Verification. — The position of the antinodes for both open and stopped pipes may be verified by the simple device of piercing the side of the pipe with small holes at the points where the antinodes are indicated for the first, second, and third overtones. These openings in the side of a narrow wooden pipe may be covered by turning a small button (Fig. 131).

The pressure at an antinode is always atmospheric. An opening made there will not then affect the pitch; but the uncovering of a hole at any other point in the pipe will be announced at once by a change of pitch. Suppose an open pipe is blown so as to give strongly the second overtone. The division of the air column is in segments with an antinode at one third the length of the pipe from either end (Fig. 130 C). If then a button be turned so as to open the pipe at either third of the length, there will be no

change in the pitch of the overtone. Similarly for the first overtone a hole may be opened at the middle of the pipe without changing the pitch.

For a stopped pipe, on the other hand, the opening for the first overtone must be made at one third the length of the pipe from the closed end, and for the second overtone one fifth or three fifths from the closed end (Fig. 129 *B* and *C*).

Fig. 131

Fig. 132

Fig. 133

238. Manometric Flames. — Manometric flames, or flames showing variations of pressure, are very suitable for exhibiting to the eye the variations of pressure produced by the voice or by a large tuning fork, or the presence of stationary undulations in an organ pipe.

A short cylinder, 3 or 4 cm. across, is divided into two chambers by a partition of goldbeater's skin or thin rubber (Fig. 132). Illuminating gas is conveyed into one of these chambers by one tube and out by another to a pin-hole burner, where it burns as a small flame. Any pure tone at the mouthpiece produces alternate compressions and rarefactions in the chamber on the left-hand side of the membrane, and these retard and aid the flow of gas to the burner. The flame changes shape and flickers, but its vibrations are too rapid to be seen directly. If now it is examined by reflection from the rotating mirrors attached to the faces of a cubical box, its image is a serrated band.

Koenig fitted three of these manometric capsules to the side of an open pipe (Fig. 133), the membrane on one side of the gas chamber forming

part of the wall of the pipe. When the pipe is blown so as to sound its fundamental tone, the middle point is a node with greatest variations of pressure, and the flame at that point is more violently agitated than at the other two. When the air blast is increased so that the fundamental gives way to the first overtone, the middle point is an antinode with no variation of pressure and the middle flame does not flicker; but the other two vibrate, and the number of tongues of flame in the image is twice as great as for the fundamental tone.

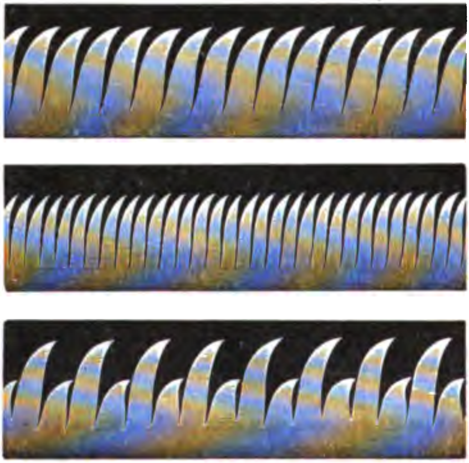


Fig. 134

The first band of Figure 134 shows the appearance of the image for the fundamental tone, the second for the first overtone; the third band may be obtained by adjusting the blast of air so that the fundamental and the first overtone are produced at the same time. The same figure may be obtained by singing into the mouth-piece of Figure 132 the vowel sound *o* on the note *B*♭, showing that this vowel sound is composed of a fundamental combined with the first overtone.

239. Kundt's Dust Tube. — The division of a resonant pipe into segments is beautifully shown by means of a glass tube

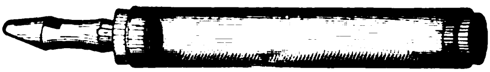


Fig. 135

about 2 cm. in diameter and 40 cm. long. One end is closed and a common whistle is attached to the other (Fig. 135). Within the tube is placed a little sifted cork dust or amorphous silica. When the whistle is blown, the powder is caught up by the moving air at the antinodes, and settles down in small circles at the nodes; at the same time between

the nodes it is divided into thin, airy segments with vertical divisions, the agitation being sufficient to support the dust in opposition to gravity. The distribution of the light powder demonstrates the presence of stationary waves due to the superposed direct and reflected systems. The subdivision changes when the blast of air is increased to give overtones.

Kundt has given to the experiment an ingenious form, designed to compare the velocity of sound in air and other media. His apparatus consists of a glass tube about a meter long and 4 or 5 cm. in diameter, closed at one end *a* (Fig. 136) by a membrane of thin sheet rubber, while at the other end is a cork piston *b* sliding freely in the tube. The inside of the tube, which must be very dry, is dusted with fine,

Fig. 136

sifted cork filings or amorphous silica. A rod of glass or metal is held by its middle point in a clamp *B*, and one end is furnished with a small disk of stiff paper, which rests against the rubber membrane at *a*.

When the rod is made to vibrate longitudinally by friction, it vibrates like an open organ pipe, giving its fundamental tone. The paper disk at *a* communicates the vibrations of the rod to the air in the tube, and when the length of the air column has been properly adjusted by means of the piston *b*, the cork dust is tossed about and gathers in small heaps, which indicate very neatly the exactness of the adjustment. The average distance l between nodes is then the half wave length of the note in the tube. The half wave length in the rod is its length L. These distances are traversed in the same time; and, therefore, if v is the velocity of sound in air at the temperature of the experiment, and V the velocity in the rod, we have the simple relation

$$\frac{l}{L} = \frac{v}{V}.$$

V. QUALITY

240. Definition of Quality. — Two of the essential characteristics of musical sounds, namely, pitch and loudness, have already been considered. There is a third important difference between musical sounds, known as *quality*. It is easily perceived that one musical note differs from another, not only in being more acute or grave, louder or softer, but also in respect to the character of the sound. We have no difficulty in distinguishing the notes of a piano from those of a violin, even though they are of the same apparent pitch and loudness. Similar differences enable us to distinguish one voice from another in speech or in song, even when somewhat modified by transmission through the telephone, or when reproduced by the phonograph. Even the untrained ear recognizes characteristic differences in music produced by different instruments of the same class.

All differences in musical notes, not assignable to pitch or loudness, are included under the term *quality*. Quality is the characteristic of a musical sound that enables one to refer it to its source.

241. Quality due to Overtones. — If an open and a stopped organ pipe, giving notes of the same pitch and loudness, are compared, a marked difference will be observed in their quality. In the former the whole series of harmonics may be present, while in the latter only those whose frequencies are odd multiples of the fundamental are possible. The sound waves in each case are the result of compounding the fundamental with the overtones present.

Pitch depends on the length of the sound wave, loudness on its amplitude, and quality on the only other physical difference between aërial waves, that is, their vibrational form. The form of a wave is the manner in which the displacements vary from point to point in the wave. It depends on the waves of higher frequency combined with the fundamental. Every change in the form of a complex sound

wave, not affecting pitch or loudness, is due to some change in the components of higher frequency than the fundamental, that is, in the overtones. Hence, *quality is to be referred to the number, the order, and the relative intensity of the overtones associated with the fundamental.*

242. Quality of Musical Sounds. — When the wave form is simple harmonic, it produces a *pure tone*, which even the musically trained ear is unable to analyze into components. When the wave form is complex and periodic, a well-trained ear can analyze the note into component pure tones, or a fundamental and harmonics.

Tuning forks and wide-mouthed stopped pipes give nearly pure tones, which are not satisfactory as musical sounds because of their dull or colorless quality. Musically, notes are described as full and rich when they consist of a fundamental accompanied by a series of moderately loud harmonics not higher than the fifth, with a frequency six times that of the fundamental. Such are the notes of open organ pipes, the French horn, and the softer tones of the human voice.

The first harmonic, with a frequency twice that of the fundamental, forms the octave; the second, with a frequency of three, gives the octave plus a fifth; the third, with a frequency of four, gives the double octave; the fourth, with a frequency of five, gives two octaves plus a third; the fifth, with a frequency of six, gives two octaves plus a fifth. The sixth overtone, with a frequency of seven times the fundamental, does not fall within the musical scale and is inharmonic. The same is true of the eighth overtone.

When a large number of overtones or upper partials are present, the notes are said to be nasal. When overtones above the fifth are quite distinct, the quality is piercing and rough on account of the dissonances introduced by the higher overtones. Such notes are said to be metallic because of their resemblance to those given by a vibrating sheet of metal.

VI. INTERFERENCE AND BEATS

243. Principle of Interference.—We have already seen that the result of combining direct and reflected sound waves in pipes is a system of stationary air waves with nodes and antinodes in fixed positions. These stationary waves are due to the superposition of the direct and reflected trains of waves. The principle of the superposition of two sets of similar waves traversing the same medium at the same time, while the resulting displacements at any point is the algebraic sum of the two displacements due to the two systems separately, is called *interference*. The name is singularly ill-chosen because each sound wave pursues its own way as if the other were not present. No observant person can fail to notice that several trains of waves may traverse the same surface of water at the same time in different directions, each train persuing its way unimpeded by the others. Even the ripples ride freely over the crests of the long waves without hindrance. "If this is interference, it is difficult to see what non-interference would be" (Lord Rayleigh).

In the phenomena of interference the attention is directed to the occurrences at definite points in the medium rather than to the passage of waves. If two sound waves of equal length and amplitude agitate the same medium at any point, and if they are in opposite phase, so that the condensation of the one falls at the same place as the rarefaction of the other, then the medium at that point is at rest and the sound is extinguished there by interference. If at any point the two waves are in the same phase, the resulting displacement is double that of either taken separately, and the sound is greatly reënforced by interference. At the nodes in an organ pipe the two waves interfering are in opposite phase and annul each other; at the antinodes, they agree in phase and reënforce each other.

244. Interference when the Frequencies are the Same.—It is well known that the intensity of the sound of a vibrating

fork held freely in the hand near the ear and turned on its stem exhibits marked variations. In four positions the sound is nearly inaudible.

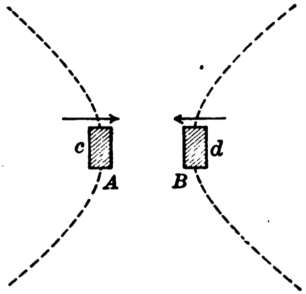


Fig. 137

Let A, B (Fig. 137) be the prongs of the fork. They vibrate with the same frequency, but in opposite directions, as indicated by the arrows. When the two branches approach each other, a condensation is produced between them, and at the same time rarefactions start from the backs at c and d. The condensations and rarefactions meet along the dotted lines of equilibrium, where partial extinction occurs. These lines lie nearly in planes passing through the axis of the fork and make angles of 45° with its face.


Fig. 138

It is easy to demonstrate that the weakening of the sound along the lines of minimum intensity is due to interference. Hold the vibrating fork over a cylindrical jar adjusted as a resonator and turn it over till a position of minimum loudness is found. In this position cover one prong with a pasteboard tube without touching (Fig. 138). The sound will be restored to nearly maximum intensity, because the paper cylinder cuts off the set of waves from the covered prong.

When an organ pipe is sounded in a large room with good

reflecting walls, regions of maximum and minimum loudness may easily be found. The variations in intensity are due to interference between the direct waves and those reflected from the walls.

Interference between waves from the same source of sound may be demonstrated by means of the so-called "duplex phonograph." In this instrument the diaphragm used to reproduce the vibrations of sound recorded on the record plate has a "horn" connected with each side (Fig. 139). When this diaphragm vibrates, it produces simultaneously a condensation on one side and a rarefaction on the other. Hence the sound waves reaching the medial line ef between the two horns a and b are always in opposite phase. As a fact of observation, when the ear of the listener is on the medial line near the horns, the intensity is noticeably less than at other points.

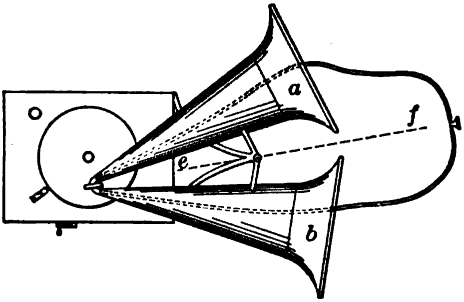


Fig. 139

The demonstration is made more complete by inserting rubber tubes in the small ends of the horns by means of tight-fitting corks, and bringing the two tubes of equal length together to a T-tube fitting the ear. The other ear should be closed. The two wave systems do not completely annul each other, but if the listener cuts off one system by pinching either tube, the intensity of the sound is increased to a surprising degree.

245. Beats.—When two sounds come from sources of slightly different period, interference gives rise to alternate swellings and subsidences in loudness, known by the term *beats*. When two tuning forks of the same pitch, mounted on resonant boxes, are sounded together, the sound is smooth as if only one fork were vibrating. Stick a small mass of wax to a prong of one of them; this load increases the moment of inertia of the fork and so increases also its periodic time of vibration. If the two forks are now sounded together, the phenomenon of beats will be very pronounced.

Mount two organ pipes of the same pitch on a bellows, and sound together. If they are open pipes, a card gradually slipped over the open end of one of them will change its pitch enough to bring out strong beats.

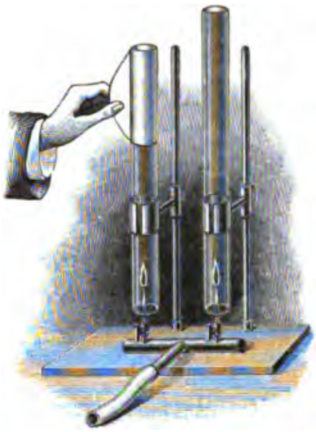


Fig. 140

With glass tubes and small jets for burning coal gas or hydrogen set up the apparatus of Figure 140. One tube should be provided with a slider for the purpose of varying its length. When the gas jets are inserted a proper distance in the long tubes and the jets are turned down slowly, both tubes will give a loud sound. The agitated flame in a sounding tube is known as a "singing flame." By adjusting the height of the slider, the two tubes may be made to give notes of the same pitch; the sound is then smooth and steady. If now the slider be moved either up or down, the sound will pulsate strongly with distinct beats.

246. Number of Beats.—Let two notes be produced by forks making, for example, one hundred and one hundred five vibrations per second respectively. Then in each second the fork of higher frequency gains five vibrations on the other; and five times during each second the two are vibrating in the same phase and five times in opposite phase. The same changing relation of phase occurs in the air transmitting the vibrations. Hence, subsidence of sound must occur five times a second, and there are five beats. Therefore, *the number of beats per second is equal to the difference between the vibration frequencies of the two notes.*

247. Beats due to Overtones.—Beats are produced not only between two notes nearly in unison, but between notes whose interval is only approximately an octave, a major third, a fifth, and so on. These are attributed to the presence of overtones associated with the fundamentals. Thus, if two

notes have frequencies n and $2n+1$, then the first overtone of the lower note will be due to $2n$ vibrations per second, and this will produce one beat per second with the slightly higher note. So also if two notes are due to $2n+1$ and $3n$ vibrations per second respectively, then the second overtone of the first will correspond to $6n+3$ vibrations, and the first overtone of the second note to $6n$ vibrations per second, and these two overtones will give three beats per second, though the interval is otherwise indistinguishable from a fifth.

Again, the interval between the fundamentals may be exact, but the frequencies of the overtones may not be exact multiples of the fundamental. Such may be the case with tuning forks, and beats are sometimes heard between their overtones of the same order.

248. Lissajous's Figures. The optical combination of the simple harmonic motions of two tuning forks, vibrating in planes at right angles to each other, was first described by Lissajous; the resulting curves are therefore known as *Lissajous's figures*. The method of obtaining them graphically has already been described in § 41. Lissajous's figures furnish a method of observing beats optically as well as by the ear; also of comparing the relative frequencies of two vibrating bodies with great precision.

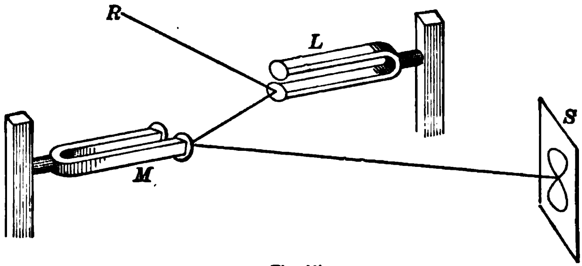


Fig. 141

A beam of light falls in succession on the plane mirrors attached to the forks L and M (Fig. 141), and is finally reflected to a screen S. If now the fork L is set vibrating,

the spot of light on the screen will appear as a vertical band, because of the persistence of impressions on the retina of the eye. Similarly, if *M* only is vibrating, the spot will form a horizontal band. If both forks are vibrating, the spot will have both harmonic motions imparted to it, and will trace a characteristic Lissajous's figure. The forks should be of low pitch, and should be kept vibrating by means of electromagnets. The same pair of forks may serve for several pairs of relative frequencies by tuning with sliding weights.

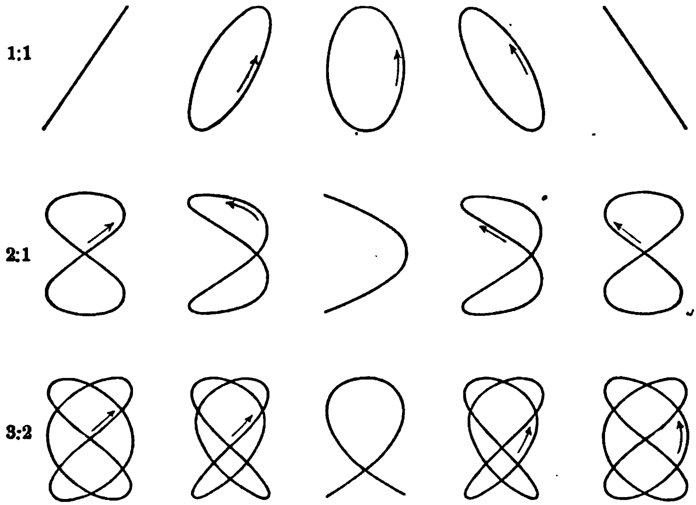


Fig. 142

If the forks are in unison, the characteristic figure is an ellipse with oblique straight lines as limiting forms (Fig. 142). If the tuning to unison is not exact but only approximate, the figure takes the successive forms of the first line from left to right, and then the same ones recur in reverse order. One beat will be heard as often as the figure goes through the whole cycle of changes from the straight line at the left over to the right and back again. During this period one of the forks gains one vibration on the other.

In Figure 142 are shown also the curves for frequencies 2 : 1 and 3 : 2. The forms corresponding to no difference of phase are in the first column; those differing by $\frac{1}{8}$ of a period, in the second column, and so on.

In all cases during the time the figure for any interval passes through a complete cycle of changes, one fork has executed one vibration in addition to the number required for the exact interval.

VII. VIBRATION OF RODS, PLATES, AND BELLS

249. Vibration of Rods. — Rods of metal, of wood, and of glass may vibrate either transversely or longitudinally. A rod fixed at one end may be made to vibrate transversely by drawing the free end aside and releasing it, or by striking it with a suitable hammer. It may be set in longitudinal vibration by clamping at the middle and stroking lengthwise with a chamois skin dusted with powdered rosin. A moist cloth is better for glass.

In the transverse vibration of rods, unlike that of strings, the force of restitution is the elasticity of flexure. The frequency of a rod with a rectangular section is independent of its width, but is directly proportional to its thickness. If two bars of the same material have the same length, while one is twice as thick as the other, the thick one will vibrate with twice the frequency of the other, whatever may be their relative width.

Examples of the use of transversely vibrating rods in musical instruments are the reeds of an accordion or harmonium, the tongue of a jew's-harp or of a music box, the reeds of reed pipes in organs, the xylophone, and the tuning fork.

The xylophone is a primitive instrument with bars free at both ends. It consists of a series of wood prisms of the proper length and thickness, supported by strings at the nodes, which are at points about one quarter of the length from each end. The prisms are adjusted to give the notes of the scale, and they are played by striking them in the middle with a little hammer having a soft, elastic face.

250. The Tuning Fork. — The tuning fork is one of the most important applications of vibrating rods free at both ends. A straight elastic bar, when sounding its fundamental, has a node at a distance from each end of about one quarter of its length. As this bar is gradually bent into the form of a tuning fork (Fig. 143), the nodes approach each other; and when the fork has a stem, the nodes are near the bottom of each branch. The two branches then vibrate in unison, while the stem has a slight up and down motion, which is transmitted to the resonant box on which the fork is mounted. The overtones are of high pitch and feeble intensity, and soon vanish, leaving a pure tone.

Fig. 143

251. Vibration of Plates. — If a plate of elastic material, such as brass or glass, be clamped at its middle in a horizontal position, and fine sand be scattered evenly over it, the sand will gather along certain definite nodal lines (Fig. 144) when the plate is thrown into vibration by drawing a bow across its edge. The sand figures increase in complexity as the number of segments becomes larger with rising pitch. These sound figures were first obtained by Chladni, and are known as Chladni's figures. The arrangement of nodal lines is

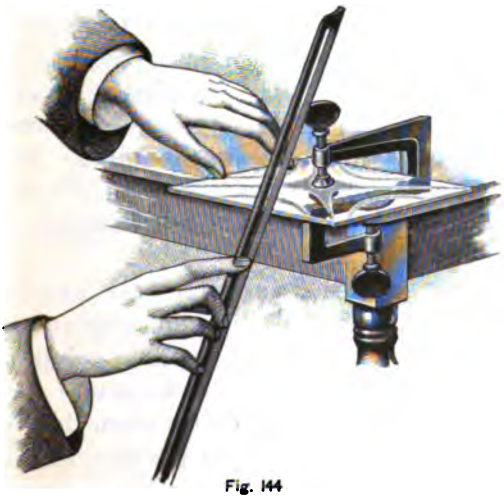


Fig. 144

determined by the position of the point bowed relative to those damped by the finger tips.

The vibration of a glass plate may be beautifully shown by supporting it close to the condensing lens of a vertical lantern, covering it with a thin film of water, and projecting the surface on the screen. The vibration of the plate produces stationary waves in the film of water covering the middle portions of the vibrating segments between the nodes. The finer stationary waves correspond to the acuter modes of vibration.

The fundamental tone of a round plate is produced by a division into four equal segments by two diameters; the first higher tone comes from a division into six segments by three diameters, the second by eight segments and four diameters. Adjacent segments, like those of strings, are always in opposite phases of motion.

The vibration of plates is illustrated among musical instruments by the cymbal, the kettledrum, and the snare drum. They have small musical value and are used solely to accentuate the rhythm. The best that can be accomplished by the tuning of a drum is to prevent its disturbing the harmony of other instruments.

252. Vibrations of Bells. — The modes of vibration of a bell are modifications of those of a cylinder. The motions of the rim are both radial and tangential. Consider a thin circular shell (Fig. 145); the forces resisting extension of this ring are very large compared with those resisting bending. We may therefore regard the perimeter as bending only and not changing in length.

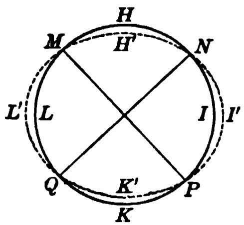


Fig. 145

For the gravest tone the number of segments is four, with nodal meridians running down the bell. The motion is entirely radial at the middle points of the segments H, I, K, L. But while the nodes are at rest radially, they are not at rest tangentially; for when the rim on one side of the node is inside its mean position, on the other side of the same node it is outside its mean position; $MH'N$ is less than MHN, and $ML'Q$ is greater than MLQ. If then there is no stretching, there must be tangential motion

of the nodes of radial vibration to allow for the variations in the length of the segments intercepted between adjacent nodes. The nodes for the radial motion are therefore antinodes for tangential motion.

The tangential motion in the rim of a vibrating bell explains the method of sounding a wineglass by drawing the wet finger round the edge. The friction produces a tangential displacement which is accompanied by a radial displacement, just as a stroke on the bell produces a radial displacement which is accompanied by a tangential displacement.

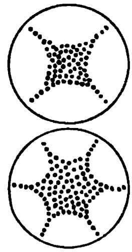


Fig. 146

If there is lack of symmetry in the figure, or of homogeneity in the rim, the nodes may slowly revolve around the rim, producing the beats often heard when a large bell is struck.

The radial motion of a bell-shaped glass vessel may be illustrated by partly filling it with methyl alcohol before bowing the edge of the glass. The alcohol is thrown in drops from the sides of the glass toward the center; these drops do not coalesce at once but form a cross (Fig. 146) for the fundamental, and a six-sided figure for the first higher tone.

Problems

1. If standard pitch is based on $a' = 435$ vibrations per second, calculate the frequency for c', both for the diatonic and the equally tempered scale.

2. Calculate the wave length of the note e', which is a third above c' of 256 vibrations per second, when the temperature is 10°.

3. A tuning fork is held over a long glass tube partly filled with ice-cold water. The most marked reënforcement of sound occurs when the column of air in the tube above the water is 32.5 cm. long. Find the frequency of the fork.

4. A wire stretched with a weight of 8 kgm. gives the note C. What must be its tension for a note a fifth higher?

5. A string 1 m. long has a frequency of 260 per second, when the tension is 28 kgm. What must be the tension so that one half the string shall have the same frequency?

6. A Kundt's tube, filled with air and thrown into vibration, divided into segments 2.5 cm. long, between successive heaps of dust. When the tube was filled with hydrogen, the segments for the same note were 9.65 cm. long. Calculate the velocity of sound in hydrogen.

7. Two telegraph sounders are in the same electric circuit which is closed five times every two seconds by the pendulum of a clock. If the listener is at one instrument, how far away must the other one be so that he may hear its nth click at the same instant that he hears the $(n+1)$th of the near one, the temperature being such that the velocity of sound is 340 m. per sec.?

8. What will be the lengths of two open organ pipes which give the notes c' and g' when the velocity of sound is 344 m. per second?

9. A string, vibrated by a large fork, is stretched with a weight of 270 gm. and divides into four segments. What must be the weight to cause it to divide into three segments with the same fork?

10. A cord vibrates synchronously with the attached fork by dividing into three segments. If it be replaced by a similar one of the same length and four times the sectional area, what relative weight will be required to cause it to divide into four segments?

LIGHT

CHAPTER VIII

NATURE AND PROPAGATION OF LIGHT

I. GENERAL PHENOMENA

253. Light and Radiation.—The word *light*, as in the case of sound, has acquired two distinct meanings: first, the common and familiar one used to designate the sensation of vision; second, the external agency acting through the eye to excite visual impressions. Nearly all relating to the first aspect of the word light lies within the field of the psychologist. Except in the study of the sensations of color, the word light in physics refers to the external cause of the sensation of luminosity, quite apart from the organ which reveals it to us. In fact, the physicist now studies light as a phenomenon of the transference of energy by wave motion. The process is known as *radiation*.

It has been found that the radiation affecting the normal eye also affects a photographic plate, a sensitive thermometer, and other appropriate detectors of radiant energy flowing from the sun, the electric light, and other similar sources. Further, the analysis of radiations from such sources reveals the fact that only a limited portion of them are capable of exciting vision. But the essential identity of these radiations with light is made evident by the fact that the various phenomena of optics may be reproduced by radiations which do not affect the eye. There is no fundamental difference between luminous and non-luminous radiations. The differences are not qualitative but quantitative; in other words, they are differences of wave length.

The limits of visible radiations are determined by the range of sensibility of the eye as a receiving instrument. It is sensitive to radiations lying within a certain rather definite range of frequency; but radiations of higher and lower frequency are not distinguished in any way from those falling within the limits of perception, except by the physical difference of wave length. It is therefore convenient and almost essential to include under light the whole range of radiations, which are alike in their fundamental properties, but which were formerly classified as luminous, heat, and actinic radiations.

254. Sources of Radiation. — The most important source of radiant energy for our planetary system is the sun. Other sources of light are solid bodies at a high temperature, such as lime at a white heat in the calcium light; glowing carbon in the carbon arc and incandescent electric lights; rare earths in the Nernst glower and the Welsbach mantle; and luminous flames, such as those of a candle, an oil lamp, a gas jet, and certain metallic electric arcs. The light from ordinary flames is radiated by glowing carbon raised to incandescence by the burning vapors. Such flames are sooty because of the innumerable solid carbon particles suspended in them, which get heated very hot and emit light.

Fig. 147

Gases do not appear to become luminous by high temperature alone, but by the passage of an electric spark or a current of electricity. The luminosity of the faint nebulæ in space still awaits an explanation. Many other subordinate but interesting sources of light are certain bodies at low temperature, which are said to give light by phos-

phorescence. A common match, when rubbed gently, gives off fumes which are faintly luminous in a dark room.

There are luminous bacteria which emit enough light to affect a sensitized photographic plate. Molisch of Prague has prepared a bacterial lamp (Fig. 147) by filling a glass jar with gelatine containing a colony of luminous bacteria. Its intensity is less than that of a candle, but bright enough to induce germinating peas to turn toward it as a source of radiant energy.

255. Nature of Light. — Light does not travel instantaneously, but it has a finite velocity. Moreover, a continuous stream of energy flows from a luminous source like the sun, for it heats bodies on which it falls. Some means must exist for the transmission of this energy. The earth receives energy from the sun; and as it takes over eight minutes for its transit across the intervening space, we are forced to seek for a vehicle by which it is conveyed.

According to our experience, there are only two ways of transferring energy; either by the projection of material bodies through space like a cannon ball, or by transmission through an intervening medium by some form of wave motion.

Newton chose the first method in his emission theory of light. He imagined the light-giving body projecting minute particles or corpuscles through space; and that they enter the eye and excite vision by impact on the retina.

The other method requires a continuous medium, called the *ether*, and energy is handed on by it from point to point by undulations. In this way energy is conveyed in sound and by water waves. A luminous body is then the source of a disturbance in the universal ether, which transfers it by undulations through space. These waves travel with the velocity of light and excite visual impressions.

Newton was aware of the wave theory, which his contemporary Huyghens had formulated; but he found an insuperable objection in the fact that water waves pass around an obstruction, and that sound shadows occur only under special

conditions. By analogy he concluded that if light is an undulation, it should also pass around bodies instead of casting a shadow.

The answer is that short water waves or ripples do not pass behind obstacles, and sounds of high pitch and short wave length cast well-defined shadows. In light the phenomena of diffraction, as we shall see later, show that light waves also pass around obstructions into the shadow.

Further, if light is transmitted by a wave motion, interference (§ 243) should apply to it as well as to sound. It was more than a century after Newton's time before Thomas Young showed by experiment the phenomena of interference in light. So great even then was the authority of Newton that Young's theories were derided by his English contemporaries, and his writings were declared to be "destitute of every species of merit." It was not until Fresnel had taken up the theory in France and had greatly extended the experimental proofs that the undulatory theory finally prevailed.

256. Light travels in Straight Lines. — *In a homogeneous medium light travels in straight lines.* A ray of light is the direction in which the radiant energy is transmitted. Light passes by an obstacle in straight lines, so that the space behind it is screened from the radiation. The sharpness of a shadow is closely connected with the excessive shortness of light waves. But even when the source of light is so small as to be considered a point, the edge of the shadow cast by an opaque body is not perfectly sharp, for the reason that points at the edge of the obstacle become new centers of disturbance by Huyghens' principle; and waves from these points overlap slightly the geometrical shadow.

When the luminous source is a point, as L (Fig. 148), the shadow will be bounded by the cone of rays ALB, tangent to the object, and will consist wholly of the *umbra*, from which the light is entirely excluded. When the luminous

body has sensible dimensions, then outside of the total shadow or umbra, and surrounding it, is a region called the

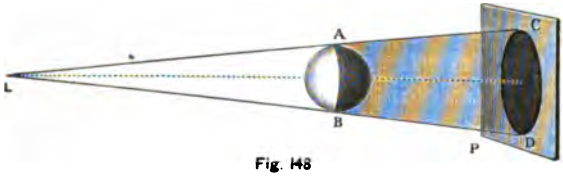


Fig. 148

partial shadow or *penumbra* (Fig. 149). This region of partial shadow forms the transition from complete obscurity to the full light. Only a portion of the luminous body is visible to an eye situated within the penumbra.

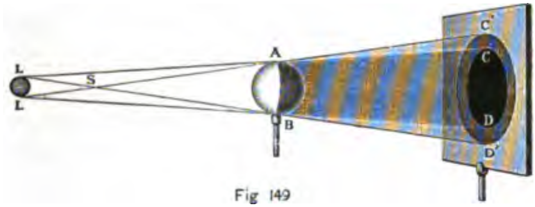


Fig 149

Solar eclipses are produced by some portion of the earth's surface passing through the shadow of the moon. The moon is smaller than the sun and its shadow is a limited cone. The apex of this shadow cone sometimes reaches the earth, when new moon occurs near one of the lunar nodes; sometimes it falls short of it, the mean length of the lunar shadow being 6700 miles less than the mean distance of the moon from the earth. If the shadow cone reaches the earth, the eclipse is total for all points lying within the umbra; within the penumbra the eclipse is partial; if only the prolongation of the shadow cone encounters the earth, the eclipse is annular for all points touched successively by the axis of the shadow.

257. Pinhole Images. — The rectilinear propagation of light is illustrated by the images produced by a small orifice of any shape, such as a pinhole. If a white screen is placed opposite a small hole in the shutter of a darkened room, an inverted picture of outside objects brightly illuminated by the sun will be formed on it in natural colors. The images

will be sharper the smaller the hole, and distant objects will be more distinctly outlined than nearer ones.

The principle is illustrated by a candle or an incandescent lamp (Fig. 150). Each point of the object, such as C, is the vertex of a cone of rays passing through the aperture and forming an enlarged image of it on the screen, as at F. The images are symmetrically placed with reference to the points emitting the light, and together they form a figure having the same outlines as the luminous or illuminated object. Since a smaller number of these images are superposed near the edges of the picture, the edges are not so bright as the other portions.

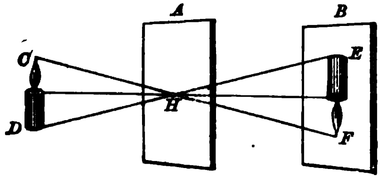


Fig. 150

If a second pinhole is made near the first one, another picture will be formed and the two will not quite coincide. Now a larger hole is the equivalent of a number of pinholes close together, and all the images overlap but do not exactly coincide. It is therefore easy to see why a very small hole is necessary for a sharp image.

The image is not only inverted, but it is *perverted*. If it is viewed from the side of the screen toward the opening, and imagined turned around in its own plane so as to make it erect, it will be found that the right side of the image corresponds to the left side of the object; that is, the image is perverted. If the screen is translucent and the image is viewed from behind, it will be inverted, but not perverted. An image in a plane mirror is perverted but not inverted.

Pinhole pictures are said to have been first observed by the great Italian artist, Leonardo da Vinci. About a century later came the "camera obscura," a small darkened chamber or box, in which was an opening larger than a pinhole and containing a converging lens. This gave more light and a sharper image. The modern photographer's camera represents the latest form of this device.

Landscape photographs and pictures of rather distant buildings of surprising softness and beauty may be made with a pinhole camera without a lens.

258. Parallax. — The apparent displacement of an object due to the real displacement of the observer is a well-known phenomenon called *parallax*. It depends on the rectilinear propagation of light. If, for example, an observer who is moving rapidly watches the moon hanging low over a hill or forest at no great distance, it will appear to travel in the same direction as the observer, while intervening objects appear to move backward.

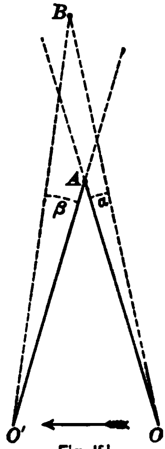


Fig. 151

When the observer is at O (Fig. 151), an object at A appears to be at the angular distance α to the left of the more distant object B; but when the observer moves to O', the object at A appears at the angular distance β to the right of B. Hence the object A is apparently displaced with respect to B through an angle $\alpha + \beta$ in a direction opposite to the motion of the observer.

If two objects are equally distant, their relative parallax vanishes. This fact serves a useful purpose in finding the position of an image formed by the objective relative to the cross thread of a telescope. When the motion of the eye right and left produces no displacement of the one with respect to the other, they coincide in position.

II. VELOCITY OF LIGHT

259. Römer's Method. — The eclipses of the inner satellite of Jupiter occur at average intervals of 48 hr. 28 min. 36 sec. These eclipses appear to take place quite suddenly, though each one is really a gradual phenomenon, and a single observation is doubtful to half a minute.

In 1676 Römer, a young Danish astronomer who was at the time an observer in the Paris observatory, announced that the observed eclipses of Jupiter's inner satellite differ systematically from the computed times. When the earth is

receding from Jupiter, the interval between two successive eclipses is longer than the mean, and the more rapid the recession the greater the excess. The reverse is true as the earth approaches Jupiter.

Let EE' (Fig. 152) represents the earth's orbit and the larger circle JJ'' the orbit of Jupiter. Then while the earth moves from E through E' to E'', or from opposition to conjunction, the eclipse intervals are longer than the mean, and the sum of all the excesses from opposition to conjunction is 16 min. 38 sec., or 998 sec. From E'' around to opposition again the intervals between eclipses are shorter than the mean, and the sum of these deficiences is again 998 sec.

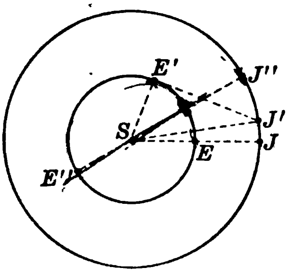


Fig. 152

Römer inferred that the speed of light is finite, so that the longer interval between two successive eclipses when the earth is receding from Jupiter is due to the additional distance which light must travel to reach the earth. This interval is greatest at E', where the earth is receding in a direct line from the planet. The sum of all the excesses is the time taken by light to travel across the earth's orbit. If the diameter be 299,000,000 km., the speed of light will be $\frac{299,000,000}{998} = 299,600$ km./sec.

Römer's original suggestion was rejected by most astronomers for more than fifty years, and was not accepted until long after his death, when Bradley's discovery of the aberration of light confirmed the correctness of Römer's views.

260. Bradley's Method.—In 1727 Bradley, afterwards Astronomer Royal of England, while attempting to measure the relative annual parallax of two stars, one of which was assumed to be much nearer than the other, discovered that every star appears to describe a small orbit about its mean

position in the period of a year. This motion is equivalent to a negative parallax; for instead of a displacement opposite to the motion of the earth, for which Bradley was looking, he discovered that a fixed star has an apparent displacement in the other direction. In the plane of the earth's orbit (the ecliptic) a star appears to move back and forth in a straight line, while a star near the pole of the ecliptic has an apparent motion in a circular orbit.

A chance observation suggested to Bradley the explanation. He noticed that the direction of a wind vane on a boat sailing on the Thames was not that of the wind, but of the resultant of the wind and a virtual head wind due to the motion of the boat. In the same way, the apparent forward displacement of a star is the change in direction due to combining the motion of the earth in its orbit with the motion of light.

Suppose the wind blowing directly against the side of a vessel moving with a velocity u, the velocity of the wind being V. The motion of the vessel produces a head wind, and this combined with the real wind causes an apparent shifting forward of the point from which the wind comes by an angle θ, the tangent of which is u/V. In the same way the velocity of light must be combined with one equal and opposite to that of the earth in its orbit in order to give the apparent direction from which the light comes, that is, to give the apparent position of a star. The phenomenon is known as the *aberration of light*.

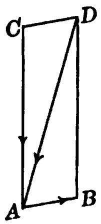


Fig. 153

Let CA (Fig. 153) represent the velocity of light V, and AB the relative magnitude and direction of the orbital velocity of the earth. Then CAD is the angle of aberration. This angle will be greatest when the motion of the earth is at right angles to the direction of the star. Then $\tan\theta = u/V$. The angle of aberration is known to be $20.445''$. Since the tangent of this angle is about $\frac{1}{10000}$, it follows that the velocity of light is about 10,000

times the earth's orbital velocity; and as the latter is very nearly 30 kilometers a second, the velocity of light is in round numbers 300,000 km./sec. = 186,400 mi./sec.

261. Fizeau's Method.—It was more than a hundred years after Bradley's discovery of the aberration of light before Fizeau in 1849 made the earliest experimental determination of the velocity of light over a limited distance on the earth's surface. His method depends on the eclipse of a source of light by means of a rapidly rotating toothed wheel.

A beam of light from a bright source S (Fig. 154) is reflected from a piece of plate glass m and focused by a lens L

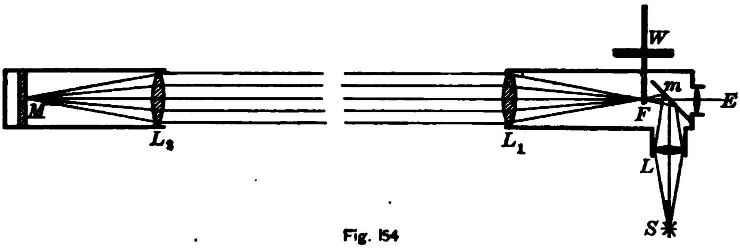


Fig. 154

on the teeth of a wheel W. The light diverging from F emerges from the lens L_1 as a parallel beam; and after traversing the intervening distance of 8633 m. to the distant lens L_2, it is focused on the mirror M, which reflects it back on its own path. The beam of light returns to the telescope and the inclined piece of plate glass; a portion of it is reflected and a portion transmitted. The latter enters the eye of the observer at E, producing the appearance of a bright star at F.

When the toothed wheel is rotated rapidly, a detached train of waves of light passes through a space between two teeth out toward M. If now the speed of light were infinite, the brightness of this artificial star would not be affected by the speed of rotation of the wheel; but if it is finite, a rate of rotation may be found such that a tooth will replace a space while the wave train travels to M and back; the returning

light will then be intercepted by a tooth and the star will be eclipsed.

What occurs in the experiment is the appearance at first of a bright star, which gradually diminishes in brightness as the speed of rotation increases, until at a definite speed it is entirely extinguished; if the speed of rotation still increases, the star reappears, reaches its former maximum of brightness, again fades away, and is eclipsed a second time when the speed of rotation is three times as great as for the first eclipse.

Fizeau found the first eclipse at a speed of 12.6 revolutions per second. The wheel contained 720 teeth, or 1440 equal divisions. Hence the time required for a tooth to take the place of a space was $\frac{1}{12.6} \times \frac{1}{1440} = \frac{1}{18144}$ sec. The double distance between the two stations was 17.266 km. The speed of light deduced from the experiment was, therefore, $17.266 \times 18{,}144 = 313{,}274$ km./sec.

In 1874 Cornu repeated Fizeau's experiment with greatly improved apparatus. His final result was 300,330 km. in air. To obtain the speed of light in the ether of outer space, this result must be multiplied by the index of refraction of air (§ 279) and this gives 300,400 km./sec.

262. Michelson's Modification of Foucault's Method. — The original experiment of Foucault was intended to compare the speed of transmission of light through water and air as a crucial test between the emission and the undulatory theory of light (§ 329). It was modified by Michelson in 1879 to make it capable of a very precise measurement of the velocity of light.

An outline of Michelson's arrangement is shown in Figure 155. S is a narrow slit, m a plane revolving mirror, L a lens of long focal length, and m' a fixed mirror. The illuminated slit becomes a source of light. It is incident on m, is reflected to m' whenever m is in a suitable position, and

forms an image of the slit at S'. When m is at rest, the light reflected from m' retraces its path to the slit S.

If, however, m has rotated through an angle θ while the light is traveling from m to m' and back, the reflected pencil will be deflected through an angle $2\,\theta$ (§ 270) and will form an image of the slit at S''.

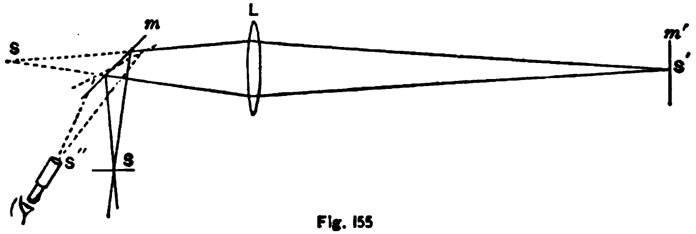


Fig. 155

If the deflection SS'' is d, the distance between S and m is r, and that between m and m' is L; and if n is the number of revolutions of the mirror m in a second and t the time required for light to pass from m to m' and back, then $2\,\theta = d/r$, $t = \theta/2\,\pi n$, and

$$V = \frac{2\,L}{t} = \frac{4\,\pi n L}{\theta} = \frac{8\,\pi n L r}{d}.$$

In Michelson's latest determination $L = 605$ m., $r = 8.58$ m., and $n = 257$. The displacement of the image of the slit was 113 mm. The final result, including all small corrections, for the velocity of light in a vacuum was 299,853 ± 50 km./sec.

By a similar method in 1882, and with a distance L between the Washington Monument and Fort Myer of 3721 m., the late Professor Newcomb obtained the value 299,860 ± 30 km./sec. It is probable that the velocity of light in a vacuum does not differ from 300,000 km./sec. by more than about one part in 3000.

III. REFLECTION OF LIGHT

263

263. Regular Reflection. — When a pencil of rays falls on a plane polished surface, the larger part of it is reflected in a definite direction (Fig. 156). The angle *IBP* between the incident ray and the normal *PB* at the point of incidence is called the *angle of incidence*, and the plane containing the two lines is the *plane of incidence;* the angle *PBR* between the reflected ray and the normal is the *angle of reflection*, and the plane containing these two lines is the *plane of reflection.*

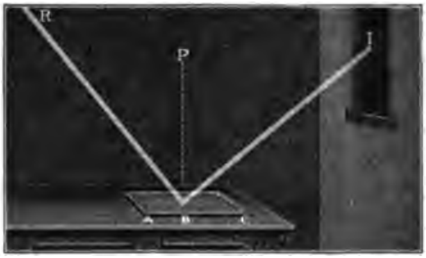


Fig. 156

The law of regular reflection is:

The angle of incidence is equal to the angle of reflection, and the planes of incidence and reflection coincide.

This law is an expression of experimental facts, and it coincides with the law of reflection deduced from the wave theory (§ 195).

264. Diffuse Reflection. — When light falls on an unpolished surface, it is reflected in all directions. Light incident on finely divided matter in suspension, such as smoke in a glass jar, is similarly reflected. The minute particles of floating smoke furnish a great many reflecting surfaces; the light is reflected in as many directions, and the result is the diffusion of the beam.

Diffuse reflection is usually selective; that is, though the incident light is white, the diffused light is colored. The rich red petal of a geranium may be illuminated by white light, but it reflects diffusely in all directions only the red.

265. Visibility. — Aëronauts have observed that at high altitudes the sky becomes black. If the atmosphere near

the surface of the earth were free from minutely divided matter reflecting the shorter waves of light, the sky on a cloudless day would appear as black as at high altitudes. What we call the sky is the pale blue light coming to us by diffuse reflection from very minute particles suspended in the atmosphere. The heavens on a moonless night are black, except for the little light of the stars and planets, even though outside the earth's shadow cone the sun is flooding space with light.

Objects are visible only when they are self-luminous or when they reflect irregularly and diffusely some of the light by which they are illuminated. Without diffuse reflection the eye would receive light from self-luminous bodies only. A beam of light admitted into a dark room is invisible if the air is dustless. It may pass through pure distilled water without illuminating it; but blow a puff of smoke across the beam, or add to the water a few drops of milk, and the light will flash out from the path of the beam.

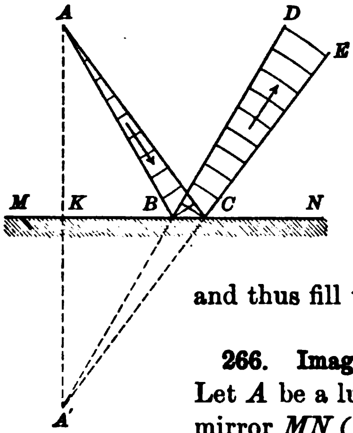


Fig. 157

It is by diffuse reflection that objects become visible. Perfect reflectors would be invisible. The trees, the ground, the grass, and particles floating in the air reflect light from the sun in all directions, and thus fill the space about us with light.

266. Image of a Point in a Plane Mirror. — Let A be a luminous point in front of a plane mirror MN (Fig. 157). The group of waves included between the limiting rays AB and AC after reflection proceed as if from A', situated on the normal AK and as far behind the reflecting surface as A is in front of it (§ 196). An eye placed at DE receives these waves as if they came directly from a source A'. The point

A' is called the image of A in the mirror MN. It is known as a *virtual image*, because the light only appears to come from it. Therefore, *the image of a point in a plane mirror is virtual, and is as far back of the mirror as the point is in front.* It may be found by drawing from the point a perpendicular to the mirror, and producing it till its length is doubled.

267. Path of the Rays to the Eye. — An image of an object is made up of the images of its points. Let AB (Fig. 158) represent an object in front of the plane mirror MN. Drop perpendiculars from points of the object to the mirror, and produce them till their length is doubled. In this manner the image of AB is found at $A'B'$. *It is virtual, erect, of the same size as the object, and perverted.*

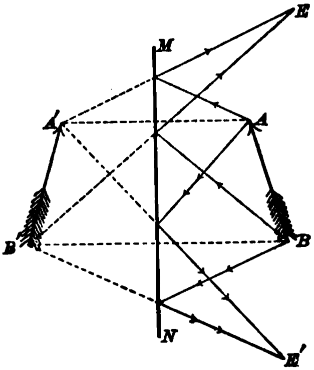


Fig. 158

It is important to observe that the image of any definite object is fixed in space, and is entirely independent of the position of the observer. The paths of the rays for the image for one observer are not the same as those for another. Let E and E' be the position of the eye for two observers. To find the path of the rays entering the eye at E, draw lines from A' and B' to E. These lines are the directions in which the light enters the eye from A' and B'. But no light comes from behind the mirror, and therefore the intersections of these lines with the mirror are the points of incidence of the rays from A and B which are reflected to E. In a similar manner the path of the rays may be traced for the position E'. The full lines in front of the mirror represent the paths of the rays from A and B, which give the images at A' and B'.

268. Multiple Reflection. — Multiple images are produced by successive reflections from two reflecting surfaces. When these surfaces are parallel, the images all lie on the normal drawn from the object to the reflecting surfaces. The double image of a bright star or planet, and the several images of a gas jet in a thick mirror, are examples of multiple reflection.

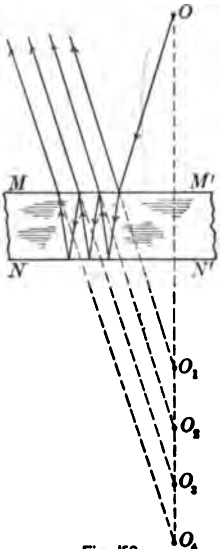


Fig. 159

Let MM' and NN' (Fig. 159) be the two parallel surfaces of a thick mirror or piece of plate glass; and let O be the object. The first image O_1 is found in the usual way. Part of the light enters the first surface, is reflected internally at the second surface, and returning to the first surface is in part internally reflected and in part transmitted, the transmitted portion forming the second image O_2, and so on. Geometrically the number of images is infinite; but on account of the loss of light by successive reflections, only a limited number are visible.

When the mirrors are inclined to each other, all the images lie on a circle, the radius of which is the distance between the object and the intersection of the planes of the two mirrors. The image in the first mirror becomes the object for the second, and this in turn is the object for a second image in the first mirror. In each case the light is incident on either mirror precisely as if it came from the next preceding virtual image in the other mirror.

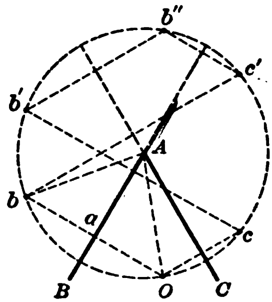


Fig. 160

Let O be a luminous point between the two inclined mirrors AB and AC (Fig. 160). Then b is the first image in

AB; and since it is in front of AC, its image in the second mirror is at c'; c' is in front of AB and has its image at b''. But b'' is behind the plane of both mirrors and there is therefore no image of it.

In the same way the images c, b', b'', may be found by first finding the image of O in AC.

By equality of triangles it is easy to show that OA, bA, $c'A$, etc., are all equal, and that therefore the object and all its images lie on the circle, the center of which is A.

269. Deviation by Successive Reflection from Two Mirrors. — The deviation of a ray of light by two reflections from a pair of plane mirrors is twice the angle between the mirrors.

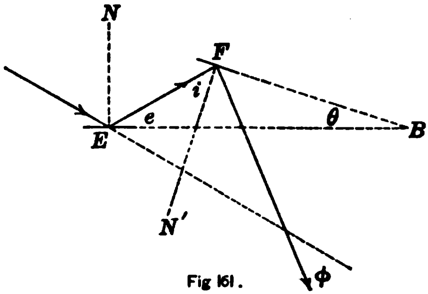


Fig 161.

Let the ray be successively reflected from the two mirrors E and F inclined at an angle θ (Fig. 161). The deviation is the angle ϕ. We have

$$\phi = 180° - 2(e + i),$$
$$\theta = 90° - (e + i).$$

Doubling the second equation and subtracting from the first, $\phi - 2\theta = 0$, or $\phi = 2\theta$.

270. Deviation due to the Rotation of a Plane Mirror. — If a plane mirror on which a pencil of rays falls be turned through an angle about an axis perpendicular to the plane of incidence, the reflected pencil will be deflected through twice the angle. Let a ray AM be incident normally on the mirror (Fig. 162); it will retrace its path after reflection. If the mirror be

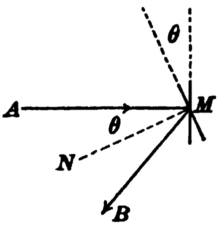


Fig. 162

now rotated through the angle θ, the normal will be turned through the same angle, and the angles of incidence and reflection will both be equal to θ. The deviation is then the angle *AMB*, or 2θ.

A plane mirror is extensively used to indicate, by the change in the direction of the reflected ray, the motion of the movable system of the instrument to which the mirror is attached. The reflecting galvanometer takes its name from the mirror which is attached to the needle system, and which indicates the slightest rotational movement.

271. Concave Spherical Mirrors.—A *spherical mirror* is one whose reflecting surface is a portion of the surface of a sphere. If the inner surface is polished for reflection, the mirror is *concave;* if the outer surface, it is *convex.* Only a small portion of a spherical surface cut off by a plane is used as a mirror. The *center* of the mirror is the center of curvature of the sphere. The middle point of the reflecting surface is the *pole* or *vertex* of the mirror, and the straight line passing through the center of curvature and the pole of the mirror is its *principal axis.*

Reflection from each element of the curved surface takes place in accordance with the fundamental law of reflection. A pencil of incident rays gives rise to a system of reflected rays, the direction of which can be geometrically determined. The position of the image in a spherical mirror can readily be determined by the application of Huyghens' principle, but the geometrical method is simpler and more generally useful.

Let *AB* (Fig. 163) be a section of a concave spherical mirror through its principal axis *AU*. *U* is a luminous point; it is required to find the equation connecting its distance from the mirror with that of its image. If a ray from *U* meets the mirror at *P*, it will be reflected across the axis at *V*, so that the radius *OP*, which is the normal at the point of incidence, bisects the angle *UPV*.

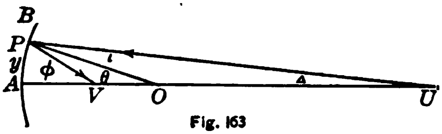


Fig. 163

Let the several angles be denoted by the letters indicated in the figure. Then from the triangles UPO and VPO,

$$\theta = i + \Delta, \qquad (a)$$

$$\theta = \phi - i. \qquad (b)$$

Adding (a) and (b), $$2\theta = \Delta + \phi. \qquad (c)$$

If now P is very near A, the angles θ, ϕ, and Δ are very small, and their tangents may be put equal to the angles themselves (§ 23). Let AU be denoted by p, AV by p', and the radius by r. Also let y denote the distance PA. Then from (c)

$$\frac{2y}{r} = \frac{y}{p} + \frac{y}{p'}.$$

Whence $$\frac{1}{p} + \frac{1}{p'} = \frac{2}{r}. \qquad (50)$$

Since y does not appear in equation (50), it follows that for a given position of the radiant point U, the distance of V from the mirror is independent of the point of incidence P. The physical interpretation is that *for small angles of incidence* all the rays from U, incident on the mirror, are reflected so as to pass through the common point V. V is a real focus because the rays after reflection actually pass through this point. V and U are called *conjugate foci*, because if V were the radiant point, the focus after reflection would be U. This relation follows from the symmetry with which p and p' enter the equation; for the values of p and p' are interchangeable, or the object and image may exchange places.

272. Focal Length. — If the luminous point is at an infinite distance, that is, if the incident rays are parallel to the axis of the mirror, p is infinite, $1/p$ is zero, and

$$\frac{1}{p'} = \frac{2}{r}, \text{ or } p' = \frac{r}{2} = f. \qquad (51)$$

The focus for rays parallel to the principal axis is the *principal focus*, and the length f, defined by equation (51), is the *focal length* of the mirror. Introducing it into equation (50),

$$\frac{1}{p}+\frac{1}{p'}=\frac{1}{f}. \tag{52}$$

When p is greater than f, $1/p$ is less than $1/f$, and $1/p'$ is positive; that is, the image is in front of the mirror and is real.

When p is less than f, $1/p$ is greater than $1/f$, and $1/p'$ is negative; that is, p' is negative, or the image is back of the mirror and is virtual.

273. Graphical Construction for Images. — *First. When the object is farther from the mirror than the principal focus.* To find the image of any point of an object, it is necessary to trace only two rays: one parallel to the principal axis, which after reflection passes through the principal focus; the other incident at the pole of the mirror, which makes the same angle with the principal axis after reflection as before. These rays are selected for convenience because they can be readily traced. They are of course not only rays, but radii of the spherical waves proceeding from the point of the object before reflection and converging toward the conjugate focal point or image after reflection.

Let AB (Fig. 164) be the object. The ray AD after reflection passes through F, and the ray AP is reflected in a direction to make the angle r equal to the angle i. The two reflected rays meet at a, which is therefore the image of A. The figure shows also the waves diverging from A and converging after reflection toward a. The point b may be found in a similar manner. The image ab is real and inverted.

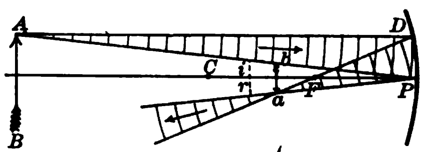


Fig. 164

Second. When the object is nearer the mirror than the principal focus. The ray AD (Fig. 165) parallel to the principal axis is reflected through the principal focus F, halfway between the pole of the mirror and its center of curvature. The ray

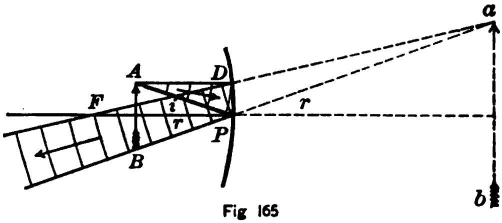


Fig 165

AP makes equal angles with the principal axis before and after reflection. The two reflected rays do not meet, but if their directions are prolonged backwards, the lines meet behind the mirror at a. This point is the virtual image of A, and the waves after reflection proceed as if from the center a. Similarly b is the conjugate focus of B and ab is the image of AB. It is erect, enlarged, and virtual.

Since corresponding portions of the object and image subtend at P equal angles i and r respectively, the size of the object and the image are proportional to their respective distances from the mirror, or

$$\frac{O}{I}=\frac{p}{p'}. \tag{53}$$

274. The Convex Mirror.—Since the center of curvature of a convex mirror is behind the reflecting surface, the radius r is negative. Then (50) becomes

$$\frac{1}{p}+\frac{1}{p'}=-\frac{2}{r}=-\frac{1}{f}.$$

But the distance of the object p is necessarily always positive; then p' must always be negative, and the formula for the convex mirror becomes

$$\frac{1}{p}-\frac{1}{p'}=-\frac{1}{f}. \tag{54}$$

The image may be found by the construction applied to the concave mirror. The ray *AD* (Fig. 166) parallel to the axis is reflected in a direction which passes through the principal focus *F*; the ray *AP* makes equal angles with the axis before and after reflection. These rays after reflection diverge as if from the point *a*, which is therefore the image of *A*. The group of rays included between *AD* and *AP* are the radii of waves diverging from *A* before reflection and from *a* after reflection. The image is erect, virtual, and smaller than the object. All possible images lie between *F* and the mirror.

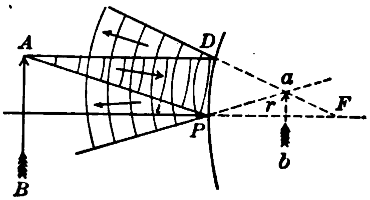


Fig 166

275. Spherical Aberration. — A semicircular sheet of tin or polished brass is placed on a sheet of white paper with its concave surface toward a candle flame (Fig. 167). The light is focused on lines called *caustic curves*. Sunlight falling on a cup partly filled with milk, or on the inner surface of a plain gold ring lying on white paper, shows by reflection the same curves.

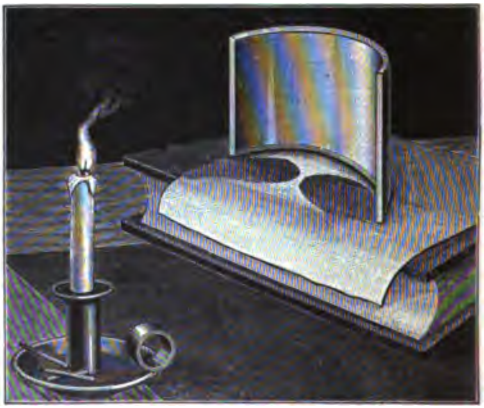


Fig. 167

Rays incident near the margin of the mirror, except when the radiant point is at the center of curvature, cross the axis at points nearer the mirror than the principal focus

F (Fig. 168). Waves reflected from a spherical mirror are not perfectly spherical, but they are normal at every point to the reflected rays. The surface enveloping the group of intersecting reflected rays is called a *caustic surface.* The cusp F of this surface is the principal focus.

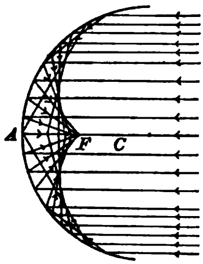


Fig. 168

Spherical aberration is the deviation from a spherical form of waves reflected from a spherical mirror. By decreasing the curvature of a mirror from the vertex outward, spherical aberration may be corrected. The reflecting surface then becomes parabolic, a form used in lighthouses, in the headlights of locomotives, and for reflecting telescopes.

IV. REFRACTION OF LIGHT

276. Refraction.—When a pencil of light passes obliquely from one transparent medium into another, it undergoes a change in direction at the bounding surface of separation between the two media. The change in the course of light in passing from one transparent medium into another is called *refraction.*

Thus, if a thin beam of bright light is admitted through a slit A in the cover of a glass jar (Fig. 169) so that it falls obliquely on the surface of water at B, its path may readily be seen and it changes direction sharply at B, the path BC making a smaller angle with the normal than that of the incident beam. This sudden deviation in the path of light in going from one medium to another was known to the ancients.

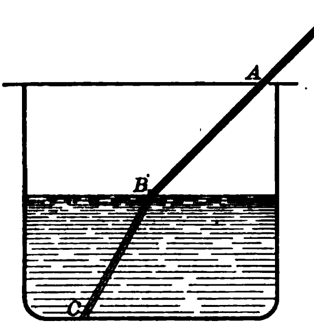


Fig. 169

277. Laws of Refraction. — Let BA denote a ray of light in air incident obliquely at A upon the surface MN (Fig. 170) of another medium, as water. AC is the refracted ray. The angle BAD between the incident ray and the normal to the surface at the point of incidence is the *angle of incidence*, i; and the angle CAE between the refracted ray and the normal is the *angle of refraction* r. The angle CAH is the *angle of deviation*. The lines BF and CG are drawn perpendicular to the normal DE. The ratio

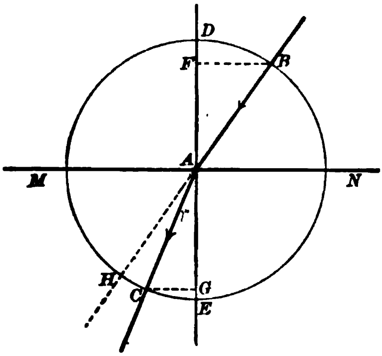


Fig. 170

$$\frac{BF}{CG} = \frac{\sin i}{\sin r} = \mu. \tag{55}$$

Snell, a Dutch mathematician, discovered, in 1621, that this ratio for the same two media is a constant, whatever the angles of incidence and refraction. The ratio μ is called the *index of refraction.*

The following are the laws of single refraction:

I. *The planes of the angles of incidence and refraction coincide.*

II. *The ratio of the sines of the angles of incidence and refraction is constant for the same two media.*

278. Velocity of Light in Different Media. — A very simple explanation of refraction was given by Huyghens by assuming that the velocity of light changes in going from one medium to another. Newton attempted an explanation by assuming an attraction between the denser medium and the corpuscles of light. By the wave theory the velocity of

light in air is greater than in water; by the emission theory it is less in air than in water.

Up to the time of Foucault, while the emission theory of light had been practically abandoned in favor of the wave theory, there had been no crucial test by experiment between the two rival assumptions relating to the velocity of light in different media. The question was definitely settled in 1850 by Foucault, who demonstrated that *the velocity of light in water is less than in air*, though he made no estimate of the ratio. In recent years Michelson measured the relative velocity of light in air and water, and found the ratio to be 1.33; also that the ratio of the velocity in air to that in carbon bisulphide is 1.76.

The retardation of light in transmission through transparent bodies is one of the fundamental facts of optics; indeed, it is probable that there is no more important phenomenon in the whole domain of radiation.

279. Explanation of Refraction by the Wave Theory. — Let AB (Fig. 171) be a plane wave incident obliquely on the smooth plane surface AC, separating air and another transparent medium. Let the velocity in air be v, and in the second medium v'. Also let t be the time required for light to traverse the distance BC. The disturbance at A has just reached the second medium, and A becomes a new center from which a spherical wave will travel in the second medium a distance AD equal to $v't$, while the disturbance in air travels from B to C, a distance equal to vt.

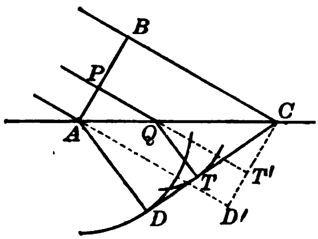


Fig. 171

In the same way the disturbance from the point P will travel to Q, which then becomes a new center; from this center the disturbance will spread into the second medium, a distance QT, such that $QT/QT' = v'/v$. All circles repre-

senting traces of such spherical waves in the second medium ultimately coalesce along CD, drawn through C tangent to the circle described about A as a center. The envelope of all these spherical waves is, by Huyghens' principle, the new wave front in the second medium.

The angles of incidence and refraction are equal to BAC and ACD respectively. They are the angles between the wave fronts in the two media and the refracting surface. Then

$$BC = AC \sin i = vt,$$

$$AD = AC \sin r = v't.$$

Whence

$$\mu = \frac{\sin i}{\sin r} = \frac{vt}{v't} = \frac{v}{v'}. \tag{56}$$

The index of refraction is therefore equal to the ratio of the velocity of light in the rarer medium to the velocity in the denser, and the constancy of the ratio of the sines has now a physical significance. The wave theory thus gives a satisfactory explanation of the laws of single refraction. The relative velocities of light in air and water, and in air and carbon bisulphide, found by Michelson, are very approximately the relative indices of refraction for both pairs of transparent media.

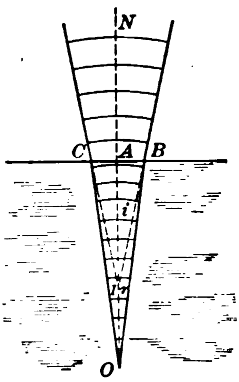


Fig. 172

280. Object Viewed in a Direction Normal to the Boundary. — Let NO (Fig. 172) be a normal to the plane bounding surface, and let the emergent group of waves between IC and IB come from the luminous point O. The center of these emergent waves is I, the image of O. The angle AIB is i, and the angle AOB is r. Then

$$AB = IB \sin i = OB \sin r.$$

Whence

$$\mu = \frac{\sin i}{\sin r} = \frac{OB}{IB}.$$

When OB is nearly normal, and the pencil of rays is only slightly divergent, OB is ultimately equal to OA, and IB to IA. IA is then equal to OA/μ. There is then no lateral displacement, but only an apparent change of distance.

The angle of the cone of rays entering the eye is limited by the size of the pupil, and is therefore small. Hence, when an object is viewed along a normal to the bounding surface, the distance of the object from the surface is μ times that of the image. For air and water μ is $\frac{4}{3}$, and for air and glass about $\frac{3}{2}$. An object in water cannot appear more than three fourths of its real depth below the surface, and one in glass not more than two thirds its actual distance from the bounding surface.

Viewed obliquely the depth of water appears less than three fourths of its actual depth. Hence the shoaling of still water when the bottom is visible.

281. The Critical Angle for Internal Reflection. — When light passes from one medium into another in which the speed is higher, its path is the reverse of the one we have hitherto considered, and a ray is deflected away from the normal in the second medium. We may still call the larger angle between the ray and the normal i and the smaller one r, even though the latter is in fact the angle of incidence. Then, since i is greater than r, there will always be a value of r for which i is 90°, and the emergent ray will just graze the boundary surface.

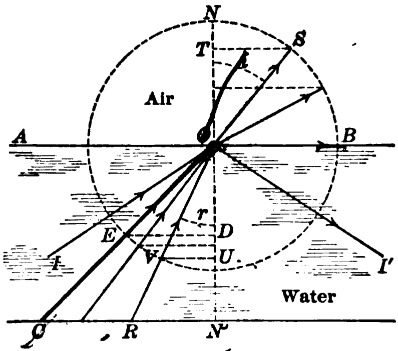


Fig. 173

In Figure 173 the ray RO makes the angle r with the normal in the optically denser medium, such as water, and the

angle i in air after refraction. The angle i increases faster than r, and the former becomes 90° when $\frac{OB}{DE}$ equals the index of refraction for the two media. The angle CON' is called *the critical angle.* For values of r exceeding the critical angle, such as ION', the ray of light can no longer emerge into the second medium, but undergoes *total internal reflection.* For angles smaller than CON', part of the light is reflected and part refracted; for larger values of r it is all reflected.

To determine the critical angle for any medium whose relative index of refraction is μ, we have

$$\mu = \frac{\sin 90°}{\sin x}. \quad \text{Whence } \sin x = \frac{1}{\mu},$$

or *the sine of the critical angle is the reciprocal of the index of refraction.*

For water the critical angle is	48° 28′
For crown glass about	41° 10′
For quartz	40° 22′
For diamond	24° 26′
For chromate of lead	19° 49′

The smaller the critical angle for a jewel with regular facets, the larger is the proportion of the light incident on it which is internally reflected. This fact explains largely the brilliancy of the diamond.

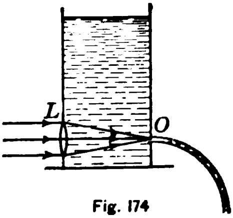


Fig. 174

The internal reflection of light is beautifully shown by focusing a beam of light by means of a lens L (Fig. 174) on the interior of a smooth jet of water at the point of issue O from the side of a tall vessel. The angle of incidence on the inside of the jet exceeds the critical angle, and the light is reflected from side to side along the stream like sound in a speaking tube. The stream is visible on account of the light diffused from fine matter suspended in the water.

282. Refraction through a Prism. — A prism for optical purposes consists of a transparent medium bounded by two plane

faces inclosing an angle, which is less than twice the critical angle for the substance. This angle A (Fig. 175) is called the *refracting angle* of the prism, and the line along which the inclined faces meet is the *refracting edge.*

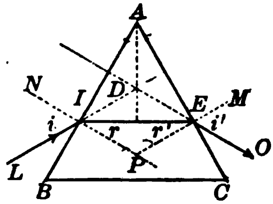


Fig. 175

Since light is refracted toward the normal when entering a denser medium, and away from it when emerging into the rarer, the path of a ray of homogeneous light through a prism may be such as $LIEO$. The plane of incidence is perpendicular to both faces of the prism, and therefore perpendicular to the refracting edge.

The deviation at the first surface of the prism is $i - r$; at the second surface, $i' - r'$. Therefore the total deviation is

$$D = i - r + i' - r' = i + i' - (r + r').$$

The angle of the prism A equals the angle between the normals at P; and since this angle is external to the triangle IPE, it equals the sum of the two interior opposite angles r and r'. Therefore $A = r + r'$.

If the path of the ray through the prism is symmetrical with respect to the two faces, *the deviation is a minimum;* then $i = i'$ and $r = r'$. It follows that $A = 2\,r$, and

$$D = 2\,i - 2\,r = 2\,i - A.$$

Hence
$$i = \frac{A + D}{2},$$

and
$$\mu = \frac{\sin i}{\sin r} = \frac{\sin \frac{1}{2}(A + D)}{\sin \frac{1}{2} A}. \qquad (57)$$

This is the formula in common use for measuring the index of refraction for any ray whose minimum deviation is D.

When the angle of the prism is very small, an approximate formula may be found by making the angles A and D

equal to their sines. Then

$$\mu = \frac{A+D}{A} = 1 + \frac{D}{A}$$

and $$D = A(\mu - 1) \qquad (58)$$

This formula is applicable to thin prisms.

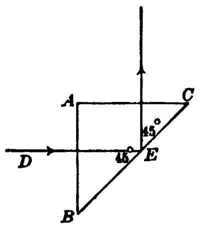


Fig. 176

For crown glass the critical angle is about 41° 10′. Hence if a prism of crown glass has as a section a right-angled isosceles triangle BAC (Fig. 176), the refracting angle will be more than twice the critical angle, and a ray DE incident normally on either face inclosing the right angle will have an internal angle of incidence greater than the critical angle, and will be totally reflected at E. Such a prism is used frequently to change the direction of a beam of light by 90°.

V. LENSES

283. Refraction at a Spherical Boundary. — A lens is a transparent body bounded by two spherical surfaces, or by one plane and one spherical surface. Before proceeding, therefore, to the subject of lenses, we must consider the case of refraction at a single spherical boundary.

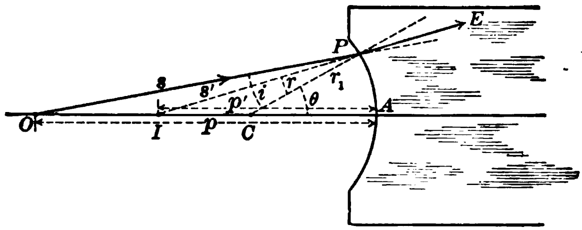


Fig. 177

Let C (Fig. 177) be the center of curvature of the concave boundary, O the point-source of light on the axis, OP the incident ray, and PE the ray after refraction; produced backward the latter meets the axis at I. I is the image of O.

Then, from the law of refraction,

$$\mu \sin r = \sin i.$$

Dividing both sides of this equation by $\sin \theta$,

$$\mu \frac{\sin r}{\sin \theta} = \frac{\sin i}{\sin \theta}. \qquad (a)$$

Remembering that the sine of an angle θ is equal to the sine of its supplement, and that the sides of a triangle are proportional to the sines of the angles opposite, we have, from the triangles CPI and CPO,

$$\frac{\sin r}{\sin \theta} = \frac{p' - r_1}{s'}, \text{ and } \frac{\sin i}{\sin \theta} = \frac{p - r_1}{s}.$$

If now the point P is near the vertex A, s' is very nearly equal to p' and s to p. Therefore equation (a) becomes

$$\mu \frac{p' - r_1}{p'} = \frac{p - r_1}{p}.$$

Dividing through by r_1 and transposing,

$$\frac{\mu}{p'} - \frac{1}{p} = \frac{\mu - 1}{r_1}. \qquad (59)$$

This equation may be considered as representing the general relation for light incident on a single boundary surface. For light *reflected* from the spherical surface, $\mu = -1$ and the formula becomes

$$\frac{1}{p} + \frac{1}{p'} = \frac{2}{r_1},$$

which agrees with equation (50) for reflection from a concave spherical mirror. If in addition r_1 is infinite, the reflecting surface is plane, $2/r_1$ is zero, and $p = -p'$, or the object and image are on opposite sides of the plane boundary and at equal distances from it (§ 266).

284. Refraction at Two Successive Surfaces. — If the ray PE (Fig. 177) is incident on a second spherical surface near the first, having a radius of curvature r_2, then, neglecting the thickness of the denser medium (the lens), the apparent distance of the object for axial rays from this surface is p'; and if q is the distance of the image, we have, from (59), the refraction being from the denser to the less dense,

$$\frac{\frac{1}{\mu}}{q} - \frac{1}{p'} = \frac{\frac{1}{\mu} - 1}{r_2}.$$

Multiply through by μ, and

$$\frac{1}{q} - \frac{\mu}{p'} = \frac{1-\mu}{r_2}. \qquad (60)$$

Adding (59) and (60)

$$\frac{1}{q} - \frac{1}{p} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \qquad (61)$$

This is the approximate formula for a thin *lens*. The distances p and q are called *conjugate focal distances*.

285. Focal Length and Principal Focus. — If the source of light O is at an infinite distance, $p = \infty$, and if we set f for the corresponding value of q, equation (61) becomes

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right). \qquad (62)$$

The distance f is called the *focal length* of the lens; it is the distance from the lens to the focus of rays parallel to the *principal axis*, the line passing through the centers of curvature of the two surfaces. This focus is called the *principal focus;* its conjugate point is at an infinite distance on the principal axis. Rays from a point-source of light at an infinite distance on the principal axis are parallel, and after

refraction they either converge toward the principal focus, or diverge as if coming from it. If the radii of curvature are equal and the index of refraction is $\frac{3}{2}$, the focal length is equal to the radius of curvature, and the principal focus on either side of the lens coincides with the center of curvature.

286. Forms of Lenses. — Lenses are in general portions of a refracting medium included between spherical surfaces. Of the six forms of Figure 178, the first three are converging

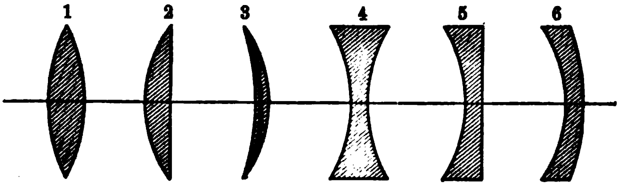


Fig. 178

lenses; they are all thicker at the center than at the edge. The last three are diverging lenses; they are thinner at the center than at the edge.

Apply formula (62) to the double convex lens 1 of Figure 178. Distances measured on the side of the lens from which the light comes are considered positive; those on the other side of the lens are negative. Then, since r_1 is negative, (62) becomes

$$-\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} + \frac{1}{r_2}\right).$$

The focal length f is therefore negative, or the principal focus for a double convex lens is on the side of the lens opposite to the source of light, and rays parallel to the principal axis are caused to converge toward the principal focus after traversing the lens.

When formula (62) is applied to 2 and 3 of Figure 178, it will be found that in these cases also f is negative, and these forms are converging.

For the double concave lens 4 of Figure 178, r_1 is positive and r_2 negative. Therefore (62) becomes

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r_1} + \frac{1}{r_2}\right),$$

and f is necessarily positive, since μ is always greater than unity. For the forms 5 and 6, f is also positive, and the principal focus lies on the same side of the lens as the source of light. The three forms, 4, 5, and 6, are therefore diverging lenses.

287. Effect of the Surrounding Medium on Focal Length. — The focal length of a lens increases as the index of refraction decreases, or f is inversely as $\mu - 1$ from (62). What then will be the effect on the focal length of a lens by immersion in another medium than air, say water?

If the index of refraction for glass is $\frac{3}{2}$ and for water $\frac{4}{3}$, then the relative index from water to glass is $\frac{9}{8}$. Therefore for a lens with equal radii of curvature, the focal length in air is equal to its radius, but its focal length in water, obtained from the relation $\frac{1}{f} = \left(\frac{9}{8} - 1\right)\frac{2}{r_1}$, is $f = 4\,r_1$. The effect of immersion in water is therefore to increase the focal length fourfold.

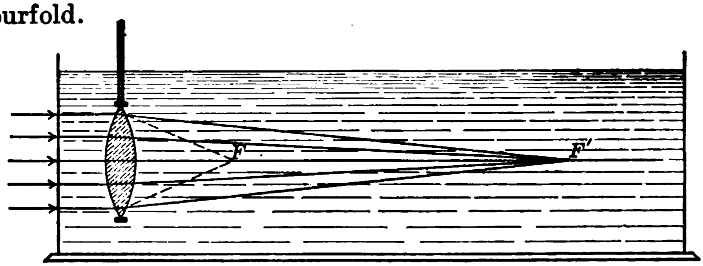


Fig. 179

Figure 179 illustrates the change in focal length by immersion in water. F is the principal focus of the lens in air, as shown by the dotted lines, while the principal focus in water is at F'.

If a converging lens is immersed in an oil in which the velocity of light is less than in glass, the converging lens becomes a diverging one. So also a hollow double convex lens filled with air and immersed in water is a diverging lens, for the speed of light in the medium inclosed in the lens is greater than in water, and the plane wave front is converted into convex wave front by the air lens. A convex air lens in oil acts like a concave glass lens in air, and a concave air lens in oil like a convex glass lens in air. All a lens can do is to change the course of light waves by impressing new curvatures on the wave fronts, and the new curvatures depend on the relative velocities of light in the lens and in the surrounding medium.

288. Universal Lens Formula.—Comparing equations (61) and (62), we may write the general formula for refraction by a lens in the form

$$\frac{1}{q}-\frac{1}{p}=\frac{1}{f}. \tag{63}$$

For diverging lenses f is always positive, with the convention here used respecting signs (§ 286). Then, since p, the distance of the point-source of light, or of the object, is necessarily positive, q must also be positive and smaller than p; otherwise $1/q-1/p$ would not be positive to conform to the formula (63). The image for diverging lenses is therefore on the same side of the lens as the object, and is virtual.

For converging lenses f is always negative. Then the general formula becomes

$$\frac{1}{q}-\frac{1}{p}=-\frac{1}{f}. \tag{64}$$

When $p>f$, $1/p<1/f$, and q in the formula must be negative, or object and image are on opposite sides of the lens, and the image is real.

But when $p<f$, $1/p>1/f$, and hence q is positive. The object and image are then on the same side of the lens and the image is virtual.

When p and q are numerically equal, the image is real, and

$$\frac{2}{p}=\frac{1}{f}, \text{ or } p=2f.$$

Object and image are then equidistant from the lens, and the distance between them is $4f$. They cannot approach nearer than this for a real image.

289. Object and Image at a Fixed Distance.—When the object, such as an incandescent lamp filament, and the screen on which the real image is received are at a fixed distance, which must be greater than four times the focal length of the converging lens, there are two positions of the lens for a well-defined image. For the first the lens L (Fig. 180) is nearer

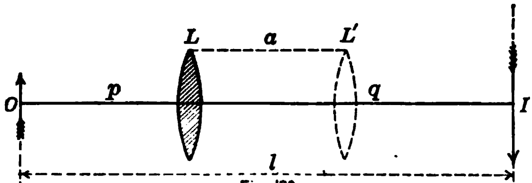


Fig. 180

the object and the image on the screen is enlarged; for the second, the two focal distances are exchanged, the lens L' is nearer the image, which is then smaller than the object.

Let l be the distance between the object and the screen, and a the distance between the two positions of the lens for distinct images. Then

$$q + p = l,$$

and

$$q - p = a.$$

Adding these equations, we have

$$q = \frac{l + a}{2}.$$

Subtracting,

$$p = \frac{l - a}{2}.$$

Therefore

$$\frac{1}{f} = \frac{2}{l + a} + \frac{2}{l - a} = \frac{4\,l}{l^2 - a^2}.$$

Whence

$$f = \frac{l^2 - a^2}{4\,l}. \qquad (65)$$

This formula furnishes a very satisfactory method of measuring the focal length of a converging lens.

290. Optical Center of a Lens. — Let C and C' (Fig. 181) be the centers of the two spherical surfaces of the lens. Draw any two *parallel* radii, as AC and BC'. Then the tangent planes at A and B are also parallel, and a ray incident at A and emerging at B passes through the lens as if through a plate with plane parallel sides, and there is no deviation; for the angles of incidence and of refraction on opposite sides of the lens are equal to each other.

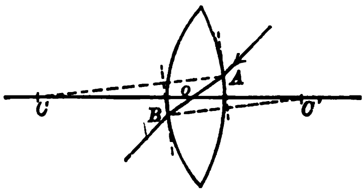


Fig. 181

AB is the path of the ray through the lens. It cuts the axis at O. Then, since ACO and $BC'O$ are similar triangles,

$$\frac{CO}{C'O}=\frac{CA}{C'B}.$$

Since the radii are constant in value, CO and $C'O$ are also constant, and O is therefore a fixed point. It is called the *optical center* of the lens. Any ray passing through the lens without deviation crosses the axis at the optical center. Any straight line passing through the optical center, except the one joining the centers of curvature, is called a *secondary axis* of the lens. In the case of a thin lens, the optical center may usually be considered as coinciding with the center of the lens.

291. Construction for the Optical Image of a Point in a Converging Lens. — To find the image of a point, or of an object consisting of a collection of points, we have only to trace two rays and find their intersection after passing through the lens. The two rays traced are the following:

1. Any ray parallel to the principal axis has after emergence a direction through the principal focus, and conversely.

C + F coincide only when $\mu = \frac{3}{2}$. See p 249. – And then not for all lenses

2. An incident ray along a secondary axis (through the optical center) emerges without deviation.

To find the real image in a converging lens, with the object AB beyond the principal focus (Fig. 182), from the

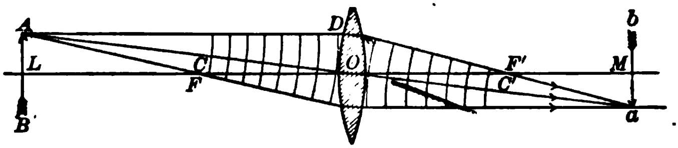


Fig. 182

point A draw the ray AD parallel to the principal axis, and continue it after emergence through the principal focus F'. Also draw a ray from A through the optical center O and find its intersection a with the first ray. Another ray through the principal focus F may be drawn; it will emerge parallel to the principal axis and will also pass through a. The point a is the optical image of A. Other points of the image ab may be found in a similar manner.

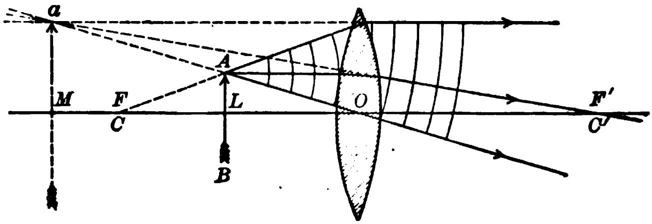


Fig. 183

For a virtual image the object AB (Fig. 183) is between the principal focus and the lens. Proceeding as before, the emerging rays now diverge as if they came from a, which is therefore the virtual image of A.

The group of waves included between the two rays traced have their wave front changed from convex to concave in the first case for a real image; and in the second case, they have their curvature decreased for a virtual image.

292. Construction for an Optical Image in a Diverging Lens.—The image formed by a diverging lens is always virtual and erect. It may be found by the same construction as for converging lenses. Thus, in Figure 184, AD, parallel

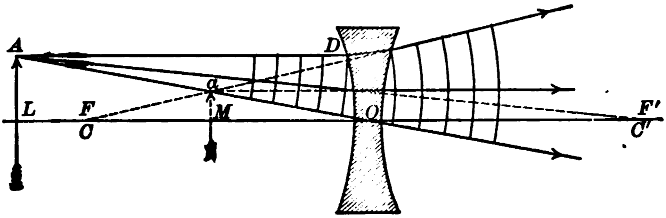


Fig. 184

to the principal axis, emerges in a direction passing through the principal focus F, and AO does not suffer deviation. The two directions intersect at a, which is the virtual image. A ray from A in the direction through the principal focus F' emerges parallel to the principal axis.

The group of incident waves have their center of curvature at A; the emergent waves, at a.

293. Spherical Aberration. — The formula for conjugate focal distances has been derived for pencils of small aperture and for thin lenses, and it has been assumed that the emergent wave is spherical. But when the aperture is large, the focal length for marginal rays is less than for axial rays, or there is noticeable *spherical aberration* by refraction. The effect is to impair the distinctness of the images formed by the lens. In a simple lens this defect is greatly reduced by the use of a diaphragm to cut off the marginal rays, leaving only the central portion of the lens effective. In compound lenses spherical aberration is corrected by combining the spherical surfaces so that their respective aberrations mutually annul each other. In large objectives for telescopes, the curvature is made to diminish toward the edge, so that all rays parallel to the axis are brought to the same focus.

If a plano-convex lens is turned with its convex surface

toward a source so distant that the incident rays are nearly or quite parallel, the path of the marginal rays through the lens is then nearly symmetrical with respect to the two surfaces, and the conditions are those for minimum deviation. The focal length for the marginal rays will then be approximately the same as for those near the axis.

VI. DISPERSION

294. Dispersion of White Light. — When a horizontal beam of sunlight is admitted into a darkened room through a narrow vertical slit, and is passed through a long-focus lens, a sharp image of the slit may be projected upon a distant white screen. If now a prism, with its refracting edge vertical, is placed near the focus of the lens, the ribbon of white light not only undergoes deviation, but it spreads out into a fan-shaped beam (Fig. 185) with its apex at the prism. This refracted beam is no longer white, but consists of a series of overlapping colored images of the slit, forming an apparently continuous band of brilliant colors, in which the hues vary continuously from red, falling nearest the white image of the slit, through all shades of orange, yellow, green, and blue to violet at the other edge of the diverging beam.

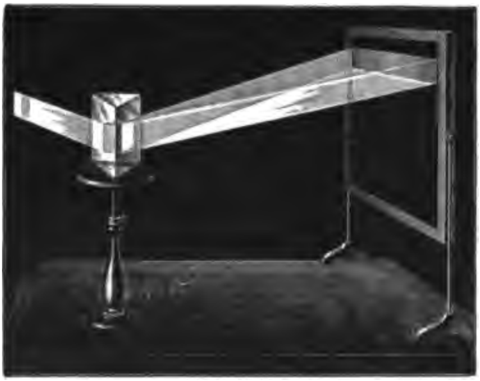


Fig. 185

This brilliant band of light, consisting of an indefinite number of colored images of the slit, is called a *solar spectrum;* and the conversion of a pencil of parallel rays of white light

into a divergent beam, in which the several colors diverge at different angles from the original direction, is called *dispersion.*

This experiment, though not original with Sir Isaac Newton, was first explained by him in 1666. He referred the colors to the complexity of white light, and concluded that the latter is to be regarded as a mixture of a series of tints, among which the "colors of the rainbow," red, orange, yellow, green, blue, indigo, and violet, were selected by him as descriptive of the series.

The deviation for red is less than for violet, and its index of refraction is therefore also less; and since the relative index is inversely as the velocity of light in the medium, it follows that red light is transmitted through the prism with greater velocity than violet. The other colors are transmitted with intermediate velocities. In fact Michelson has shown by direct measurement that the velocity of red light is 1.4 per cent greater in water and 2.5 per cent greater in carbon bisulphide than that of blue light. Dispersion is therefore due to the unequal retardation in the speed of transmission of the different colors through transparent media. Violet suffers a greater retardation, or travels more slowly, than red when it enters an optically denser medium. Measurements of wave length show that the undulations corresponding to extreme violet are the shortest of all those lying within the visible spectrum. Physically the differences in spectral colors are differences of wave length, and the short waves suffer greater diminution of velocity in a dense transparent body than long waves. In the ether of space waves of all lengths travel with the same velocity.

295. Synthesis of White Light. — Newton discovered that none of the component colors of solar light undergoes further resolution or change in kind by transmission through a second prism with its refracting edge turned in the same direction as that of the first. But when the second prism is exactly like the first and is reversed in position (Fig. 186),

there is formed a colorless image on the screen. In this experiment there is actual synthesis of white light from the spectral colors. The incident beam S is resolved into its spectral colors by the first prism, and these are recombined into an emergent beam of white light E by the second prism.

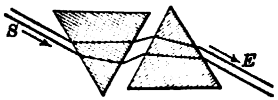


Fig. 186

A beam of white light from any source, such as the electric arc, gives substantially the same succession of colors by dispersion as those obtained from sunlight.

296. Fraunhofer Lines. —If the slit used to obtain the solar spectrum be made narrower, the colored images of it will be narrower and the overlapping of the spectral images will be diminished. Wollaston was the first to use a narrow slit, in 1802, and thus to secure an approximately *pure spectrum*, in which the colored spectral images do not much overlap. Wollaston's method was to view a narrow slit at a distance of 10 or 12 feet through a prism held near the eye. He obtained in this way virtual images of the slit and was able to detect certain dark lines or spaces crossing the solar spectrum, where the corresponding images were absent. Later Fraunhofer with better optical means counted about 750 of these dark lines, and marked the place of 354 of them on his map. They have since been called *Fraunhofer lines*.

Fraunhofer designated the most conspicuous dark lines by the letters A, B, C, D, E, F, G, H (Fig. 187). The A lines are found in the extreme red, the D lines in the yellow, and the H lines at the limit of the violet end of the visible spectrum.

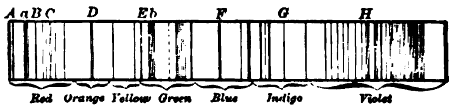


Fig. 187

Photography reveals the existence of lines like those of Fraunhofer in the ultra-violet portion of the solar spectrum

beyond the ordinary visible limit (Fig. 188). All of them indicate the absence of certain colors, or radiations of definite wave length, in the light of the sun. They have been the means of making important discoveries relating to the con-

Fig. 188

stitution and physical condition of the sun and of the stars. They serve also as a convenient means of reference for colors. Thus, when reference is made to any particular color, as D light, for example, light corresponding in wave length to the dark line D in the yellow of the solar spectrum is meant.

297. Refractive Indices and Dispersive Power. — In the following table the indices of refraction of several substances are given for the most conspicuous Fraunhofer lines :

	A	B	C	D	E	F	G	H
Crown glass	1.5089	1.5109	1.5119	1.5146	1.5180	1.5210	1.5266	1.5314
Flint glass .	1.6391	1.6429	1.6449	1.6504	1.6576	1.6642	1.6770	1.6886
Water . .	1.3284	1.3300	1.3307	1.3324	1.3347	1.3366	1.3402	1.3431
Carbon disulphide .	1.6142	1.6207	1.6240	1.6333	1.6465	1.6584	1.6836	1.7090

The angle between the divergent rays for any two colors, produced by a prism from parallel rays of white light, is called the *dispersion* for these colors. These colors undergo not only dispersion, but deviation as well, and it is of great importance in optics to know whether the dispersion always bears the same relation to the mean deviation. The ratio of the dispersion for any two colors to the deviation of the mean between them is called the *dispersive power* of the substance of which the prism is made.

If μ_1 and μ_2, D_1 and D_2, are the indices of refraction and the deviations respectively for any two colors, represented by the two Fraunhofer lines A and H for example, and μ and D

are the same quantities for some intermediate color, such as the D line, then (58)

$$D_1 = A(\mu_1 - 1),$$
$$D_2 = A(\mu_2 - 1),$$
$$D = A(\mu - 1).$$

The dispersion, that is, the angular separation of these two extreme colors, is

$$D_2 - D_1 = A(\mu_2 - \mu_1).$$

The following are the angular dispersions for the A and H lines in terms of the refracting angle A of the prism, taken from the table above:

Crown glass	0.0225 A
Flint glass	0.0495 A
Water	0.0147 A
Carbon disulphide	0.0948 A

Hence a hollow prism filled with carbon disulphide will give a spectrum 6.45 times as long as if it were filled with water.

The ***dispersive power*** expresses the property of dispersion possessed by a substance irrespective of the refracting angle A of the prism, so long as this angle is small. It is the ratio

$$d = \frac{D_2 - D_1}{D} = \frac{A(\mu_2 - \mu_1)}{A(\mu - 1)} = \frac{\mu_2 - \mu_1}{\mu - 1} = \frac{\Delta\mu}{\mu - 1}.$$

This ratio between the difference of deviations of two selected colors or lines of the spectrum and the mean deviation is constant for the same substance, so long as the refracting angle of the prism is small, but it is different for different substances. Thus for crown glass, and for the A, H, and D lines, the dispersive power is 0.0437, while for carbon disulphide it is 0.1497. *For the same mean deviation*, therefore, a hollow prism filled with carbon disulphide will give a spectrum 3.4 times as long as the one produced by a prism of crown glass.

298. Chromatic Aberration. — Since the homogeneous colors of white light have different indices of refraction, it follows from the formula

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{r} - \frac{1}{r'}\right),$$

that a single lens has different focal lengths for different colors, and that f is less as μ is greater. Hence violet light comes to a focus nearer the lens than red. Thus, in Figure 189, v is the principal focus for violet rays, and r for red. The other colors have their respective foci between these two. If, therefore, a screen be placed at or near v, as at x, the image will be bordered with red; if at y, near the focus r, it will be fringed with violet. The image with least color will be obtained by placing the screen midway between the two foci v and r, where the refracted beam has the smallest cross section; but it is impossible to get a colorless image with a single lens. This confusion of colored images is called *chromatic aberration.*

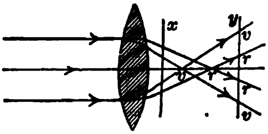


Fig. 189

299. Condition of Achromatism. — It will be apparent from § 297 that by varying the refracting angles of two thin prisms of different materials, such as crown and flint glass, and by combining them with their sharp edges turned in opposite directions, it is possible to secure *deviation without dispersion.*

The expression for dispersion is $A\,\Delta\mu$. Let the dispersion for a second prism be denoted by $A'\Delta\mu'$. Then if the image of the slit is to be colorless, the dispersion of the colors to be reunited in one direction must equal their dispersion in the other direction; that is, the angles of the prisms must be so chosen that the dispersion shall be the same for both. Then

$$A\Delta\mu = A'\Delta\mu', \text{or } \frac{A}{A'} = \frac{\Delta\mu'}{\Delta\mu}. \qquad (66)$$

This equation expresses the condition for a colorless image, or achromatism, for two prisms. Expressed in words it is, the refracting angles of the prisms must be inversely proportional to the differences between their indices of refraction for the pair of selected colors to be reunited. Strictly only the two colors selected are perfectly reunited, for the reason

that, while their angular separation is the same for the two prisms, the intermediate colors do not occupy precisely the same relative spaces in the two spectra.

To unite B and H lines near the two ends of the spectrum, the ratio of the refracting angles for crown and flint glass may be found as follows:

For soft crown glass and the B and H lines

$$\Delta\mu = 1.5314 - 1.5109 = 0.0205.$$

For dense flint glass and the same lines

$$\Delta\mu' = 1.6886 - 1.6429 = 0.0457.$$

Whence $$\frac{A}{A'} = \frac{0.0457}{0.0205} = 2.23.$$

But since these two prisms have different *dispersive powers* and produce the *same dispersion*, the *deviations* cannot be the same; the resultant deviation is the difference in deviations and is in the direction of that due to the prism of smaller dispersive power. In Figure 190, C is the crown glass prism and F the flint glass. The dispersion due to C is neutralized by the dispersion of F in the opposite direction, but there is an outstanding difference in the deviations; that is, the emergent pencil is colorless and deviates from the direction of the incident pencil.

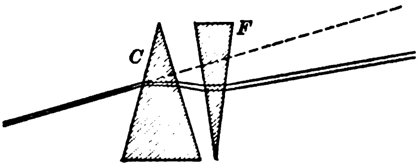


Fig. 190

In a similar way a converging lens of crown glass and a diverging lens of flint glass (Fig. 191) may be combined so as to give a nearly colorless image, and the pair may still have any desired focal length. For this purpose the focal lengths of the two lenses must be proportional to the dispersive powers of the two kinds of glass. The crown glass lens has the shorter focal length. Such a combination is called an *achromatic lens*.

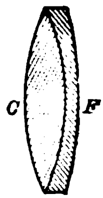


Fig. 191

300. The Rainbow. — The rainbow is the most conspicuous example in nature of dispersion on a large scale. It is due to sunlight refracted and internally reflected by spherical drops of rain. In general, the light thus refracted and reflected is dispersed in every direction, and not enough is received by the pupil of the eye to produce an impression of

color except at a particular angle. This angle is the one at which an emergent pencil of *parallel rays* leaves the drop. For parallel rays the decrease in intensity with distance is small and is dependent on absorption in the intervening medium. Since the refrangibility is different for different wave lengths, the angle at which parallel rays emerge is not the same for red as for violet. Hence the dispersion and the spectral colors of the bow.

Let O (Fig. 192) be the center of a raindrop. It was demonstrated by Descartes that there is an angle of incidence on the drop which gives the least deviation from the original direction of the parallel rays from the sun. For red this angle AOB is about 59°. If the angle of incidence is either less or greater than this, the deviation is greater. Now near a minimum (or a maximum) value the change is very slow. Hence a small parallel pencil of incident rays becomes at this angle a parallel pencil of emergent rays, and the intensity is sufficient to produce a visual impression.

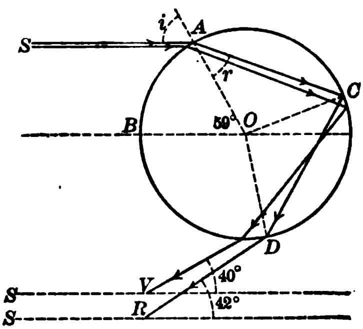


Fig. 192

In the inner or primary bow, for red, the emergent pencil makes an angle of 42° with the line drawn through the sun and the eye of the observer; for violet, the index of refraction of which is larger, the corresponding angle is 40°. In the primary bow, therefore, the red is on the outside and the violet on the inside of the circle. Spherical drops at an angular distance of 42° from the axis through the eye and the sun, and in any plane through this axis, will therefore send red light to the eye, and the bow is accordingly circular.

The primary bow is the inner and brighter one; the

secondary bow is much fainter because the light forming it suffers two internal reflections, neither of which is total. Moreover, the order of colors, as compared with those of the primary, is reversed (Fig. 193). An observer at *E*, with his back to the sun, receives red light from drops at an angular distance of 42° from the axis *SC*, and violet from those at an angular distance of 40°. For the secondary bow, the angular distance is 51° for red and 54° for violet.

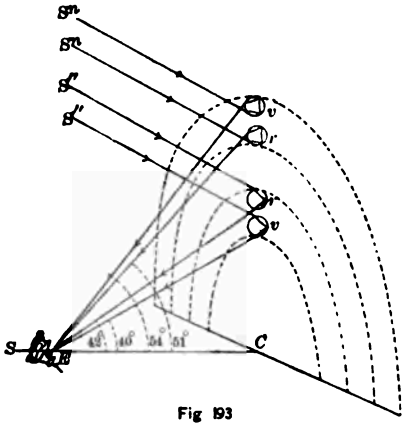


Fig 193

Artificial rainbows may be made by covering the opening of a porte-lumière with a large sheet of white cardboard, in which is a circular hole 3.75 cm. in diameter, and causing a horizontal beam of sunlight admitted through the hole to fall on a spherical glass bulb 4 cm. in diameter and filled with water. Two circular spectra, resembling rainbows, will be reflected back to the cardboard; the space between the inner and the outer bow will be quite dark.

Problems

1. An illuminated vertical object 6 ft. long is at a distance of 12 ft. from a shutter in which there is a minute hole. Inside is a vertical screen 4 ft. from the small aperture. How long is the image of the external illuminated object?

2. If a man is 5 ft. 11 in. tall, what is the length of the shortest plane mirror in which he can see his full-length image? Does the vertical distance of his eye below the level of the top of his head make any difference in the length of the mirror?

3. The ceiling of a room 20 × 30 ft. is 10 ft. 6 in. high. An observer stands with his eye in the center of the room. What is the least height of a plane mirror on one wall to enable him to see the image of the opposite wall from floor to ceiling?

4. If a plane mirror is moved parallel to itself directly away from an object in front of it, how much faster does the image move than the mirror?

5. The radius of curvature of a concave spherical mirror is 80 cm. If a pencil of light diverge from a point on its principal axis 90 cm. in front of it, at what point will it focus?

6. An object is 60 cm. in front of a concave spherical mirror and its image 20 cm. in front. What is the principal focal length of the mirror?

7. An object 12 cm. long is placed symmetrically on the axis of a convex spherical mirror at a distance of 24 cm. from it; the image is 4 cm. long. What is the focal length of the mirror?

8. A candle flame is placed at a distance of 80 cm. from a concave mirror made from a sphere of 30 cm. diameter. Find the position of the image. Is it erect or inverted?

9. If the image of an incandescent lamp is four times as far from a concave mirror as the lamp itself, what are the relative dimensions of object and image?

10. The diameter of the sun is $\frac{1}{110}$ of its distance from the earth. How many inches in diameter will be the sun's image in a concave mirror whose focal length is 48 ft.?

11. An incandescent lamp is moved from a position 25 cm. to one 30 cm. in front of a concave mirror whose focal length is 20 cm. How far is its image shifted?

12. Two plane mirrors face each other and are parallel. An object between them is distant 15 cm. from one and 20 cm. from the other. Find the positions of the first two images in each mirror.

13. An object 5 cm. long in front of a converging lens has an image 20 cm. long on a screen 100 cm. from the lens. What is the focal length of the lens?

14. At what distance from a converging lens of 30 cm. focal length must an object be placed so that the linear dimensions of the image will be four times those of the object?

15. The focal length of a converging lens is 20 cm., and the distance between the object and the screen 100 cm. Where must the lens be placed to give a sharp image?

16. The focal length of a glass lens in air is 20 cm. If the indices of refraction of glass and water are $\frac{3}{2}$ and $\frac{4}{3}$, respectively, what is the focal length of the lens in water?

17. A converging lens is placed at a distance of 40 cm. from a luminous object and forms an image of it on a screen. When the lens is moved 50 cm. nearer the screen, another and smaller image is formed. What is the focal length of the lens?

18. A glass beaker is filled with water to a depth of 5.2 cm. A cross scratched on the bottom of the beaker inside appears when viewed through the water with a microscope to be 1.29 cm. above the bottom. Calculate the index of refraction of water.

19. Find the dispersive power of crown glass and flint glass for the lines A, H, and D.

20. Two parallel walls are 16 ft. apart. Where must a converging lens of 3 ft. focal length be placed in order to project on one wall the image of an object on the opposite wall? What is the magnification?

21. A copper cent is 19 mm. in diameter and a silver half dollar 30.4 mm. At what distance from a converging lens, whose focal length is 10 cm., must a cent be placed so that its image shall be just the size of a silver half dollar?

22. If the velocity of light in air is 300,000 km. per second, what is it in water, in flint glass, and in diamond (indices of refraction, 1.332, 1.65, and 2.47, respectively)?

23. Compute the critical angle for diamond, if the index of refraction is 2.47.

24. If the apparent depth of a fish below the surface of still water is 3 ft., what is its real distance?

25. A plano-convex lens, radius of curvature 15 cm., index of refraction 1.5, has at the center of the plane side a cross scratched. The thickness of the lens at the middle is 2 cm. If it lies with the flat side on white paper, how far down in the glass will the cross appear to an observer looking down along the axis?

CHAPTER IX

LIGHT AS A WAVE MOTION

I. INTERFERENCE AND DIFFRACTION

301. Colors of Thin Films. — Up to this point it has been assumed that light is a wave motion in a hypothetical medium called the ether. We are now to consider phenomena which do not appear to admit of explanation on any other theory. The phenomena of diffraction in light not only answer Newton's objection to the wave theory, but as developed by Young and Fresnel become the strongest evidence in its favor.

If light consists of waves in an appropriate medium, it may be anticipated that interference phenomena will be observed similar to those in sound (§§ 243–246). In fact thin films of transparent substances, such as a soap bubble, a layer of oil on water, a coating of varnish on white cardboard, a film of oxide on polished or molten metal, and a thin layer of air between good reflecting surfaces, all exhibit beautiful colors by reflection in white light; and in light of one color, alternate bright and dark bands or rings analogous to beats in sound. These iridescent colors are the residuals of white light after waves of definite length have been cut out by interference between the two wave systems reflected from the parallel surfaces of the film.

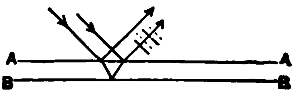


Fig. 194

Let AA and BB (Fig. 194) be the two surfaces of the film. Light incident on the first surface is in part reflected and in part transmitted. The transmitted portion is in part reflected

from the second surface, and it emerges from the first along with light externally reflected at the point of emergence.

If now the additional path through the film traveled by the internally reflected light is a whole wave length, then the two systems will be in step unless a phase difference is introduced in some other way. If the difference of phase depends only on difference in path traveled, then when this latter difference vanishes with a film of infinitesimal thickness, the two pencils of light should mutually support each other, and the illumination should be a maximum. But the fact is that when a film is made as thin as possible, it becomes black. The light is extinguished by interference. This happens for all wave lengths.

The retardation of the one system with respect to the other by half a wave length occurs at the boundaries of the thin film. One of the two interfering systems loses half an undulation relative to the other because one is reflected in the rare medium next to the dense, and the other in the dense medium next to the rare. This change in phase is analogous to that of a sound wave reflected from the open end of an organ pipe. When a sound wave is reflected from the closed end of a pipe, there is a change in sign of the motion, but a condensation is reflected as a condensation; when the reflection is from the end of an open pipe, there is no change in sign of the motion, but the condensation changes sign, and is reflected as a rarefaction. So two pencils of light reflected under the corresponding opposite conditions have impressed upon them a difference of phase equal to half a period. The constraint of the ether in dense matter and its relative freedom in space give rise to an effect similar to the phase difference impressed upon two sound impulses when reflected respectively from the closed end of a pipe and from the free air at the open end.

When the thickness of the film and the angle of incidence are such that one pencil falls behind the other in transmission by a whole number of wave lengths, interference takes

place with extinction of light, since a phase difference of half a period must be added because the reflection at the two surfaces of the film is under opposite conditions.

With white light the extinction of one spectral color by interference leaves colored fringes. Further, the thickness of film that impresses upon the internally reflected pencil a retardation of one wave length for violet produces a retardation of only about half a wave length for red (§ 307). Therefore extinction of both colors cannot occur at the same part of the film. If the violet is cut out, the red remains. The case is similar with the intermediate colors. The reflected light is, therefore, fringed with color.

302. Newton's Rings. — Sir Isaac Newton ingeniously determined the relation between the colors given by a thin film of air and its thickness by means of a film between two pieces of glass, one flat and the other slightly curved. The curved surface is a convex lens of great radius of curvature. If this radius and the distance of any point from the point of contact *C* (Fig. 195) are measured, the thickness of the film at the point is readily calculated.

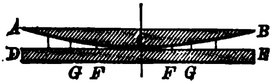


Fig. 195


Fig 196

Between the lens and the plate there is a very thin film of air, increasing in thickness outward from the point of contact. At the contact there is a dark central spot, and around this are concentric colored rings or "fringes" (Fig. 196). When the light is incident normally, the rings are circular; when the incident is oblique, the rings are elliptical. If the two pieces of glass are forcibly pressed together, the colored rings expand and the distortion of the glass is sufficient to distort the rings also.

When the illumination is by red light, the rings are dark and bright in succession, and their diameter is larger than if the light were blue, indicating that the waves of red light are longer than those of blue. A determination of the wave length of any spectral color may be made by measuring the diameter of a ring of that color and the curvature of the lens. This method is inferior to others now in common use. The wave length of yellow or *D* light is about 0.000059 cm. or 1/43000 in.

303. Transmission of Waves through Narrow Apertures. — When a beam of sunlight is admitted into a darkened room through a very narrow slit, and is received on a white screen at some distance, there will be a central band of white light on the screen in the direct path of the beam, bordered with colored fringes. Through so narrow an opening light not only passes as a definite pencil, but it also diverges from all points of the opening as new centers of disturbance (Huyghens). The colored bands are due to interference of the diverging secondary waves, and they are called *diffraction fringes.*

Diffraction may be explained as a phenomenon of wave motion and applicable to both sound and light. Let *ab* (Fig. 197) be the width of a narrow aperture, through which

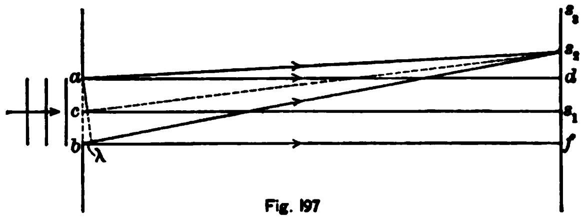


Fig. 197

come plane waves of one wave length λ. When such a wave reaches *ab*, all points along the line *ab* are in the same phase of vibration. The lines *ad* and *bf* are the boundaries of the geometrical pencil or wave train.

Now assume that the relation between the wave length and

the width ab is such that the point s_2, for which the distance bs_2 is exactly one wave length greater than as_2, lies outside the geometrical beam df. Then the secondary waves from a and b will arrive at s_2 in the same phase, since the retardation of the waves from b as compared with those from a is exactly one wave length. But the waves from a and c (the middle point of ab) will arrive at s_2 in opposite phase, since cs_2 exceeds as_2 by half of λ. They will therefore annul each other.

Similarly for every point between a and c there is another between c and b so situated that the difference in their distance from s_2 is half a wave length. Hence all the disturbances coming from ac completely annul at s_2 the disturbances coming from the other half cb of the slit.

Next consider a point s_4, the distance of which from b exceeds that from a by two wave lengths. Then the distance from c to s_4 is one wave length greater than the distance of a from s_4, and the conditions for this half of the slit are those for complete extinction of the disturbance at s_4 as just explained. The same is true for the other half cb. Therefore there is complete neutralization of all the disturbances of wave length λ arriving at s_4 from the slit ab. The same is true for any point, the distances of which from the edges of the slit ab differ by any even number of times $\lambda/2$.

If the distances of a and b from the point s_3 differ by three halves λ, then the slit may be divided into three equal parts, and the secondary waves from two adjacent ones of these will interfere at s_3 in the manner explained, while those from the third part alone are effective. Another point s_5 is so situated that its distance from b exceeds its distance from a by five half wave lengths. Secondary waves from four of the five equal divisions of the slit then interfere in pairs at s_5, leaving only one fifth of the slit to send effective waves to s_5. These equal divisions of the slit are called *half period elements*.

In general if $bs - as$ on either side of s_1 is an even number

of half wave lengths, there is an even number of half period elements in *ab*, the secondary waves from which interfere in pairs at s; but if $bs - as$ is an odd number of half wave lengths, then *ab* may be divided into an odd number of half period elements, and the surviving secondary waves from only one of these arrive at s. With monochromatic light, bright and dark bands alternate on either side of s_1. The position of these bands is given by the equation

$$bs - as = n\frac{\lambda}{2}, \tag{67}$$

where n is even for the dark and odd for the bright bands.

When the slit is illuminated by white light, in which λ is different for the several spectral colors, extinction for these colors will occur at different distances from s_1, and therefore the fringes on the screen are colored.

304. Conditions for Diffraction Phenomena. — The resulting disturbance is a minimum at s_2, s_4, etc., and a maximum at s_3, s_5, etc. It is obvious that the successive distances between maxima decrease outward; moreover, these maxima themselves decrease rapidly in intensity, for s_3 receives secondary waves from one third of the slit as a half period element, s_5 from only one fifth, etc. Hence at no great distance on either side of s_1 there is a region which no disturbance reaches. The width of the area within which diffraction phenomena may be observed is determined by the relation between the wave length of the disturbance and the width of the opening *ab*. It will readily be seen that if the width of the opening is many times the wave length, the points s_3, s_5, etc., for the chief maxima may all lie within the geometrical beam between *ad* and *bf* (Fig. 197). When, therefore, the wave length is very small in comparison with the width of the opening, the wave motion is propagated in straight lines through the opening and does not diverge appreciably into the geometrical shadow. This is the condition for sharp shadows for both sound and light.

Again, if the opening through which the waves come is less than a wave length in width, there is no point such as a_2, the distance of which from one edge of the aperture is a whole wave length greater than from the other edge. Under this condition there is no region of complete interference, and the geometrical outline of the shadow disappears.

If these conclusions are applied to waves of light and of sound, it will at once be apparent why ordinarily sharp shadows are observed for light only and not for sound. Since the mean wave length for light is only about 0.00005 cm., the diffraction bands for the usual openings all lie so near the edge of the geometrical beam that they are indistinguishable from it. They can be observed only when the opening is very narrow.

On the other hand, sound waves of mean pitch have a length of about 125 cm., or 4 feet. Hence ordinary openings, such as a window, have a width comparable with the wave length, and the secondary waves diverge in all directions beyond the opening, or there is no well-defined sound shadow. The particular phenomena of diffraction in sound were not recognized until after their discovery in the case of light. (Consult § 211.)

305. The Diffraction Grating. — The phenomena of diffraction with a single slit are not readily observed. They are best shown by a *diffraction grating*, which was devised by Fraunhofer nearly a century ago.

A diffraction grating consists of a surface containing a very large number of equidistant parallel lines for the transmission or reflection of light. A transmission grating is made by cutting the lines on a plate of glass with a diamond point by means of a dividing engine. The light passes through the transparent spaces between the ruled lines. A good grating has 1000 to 5000 lines to the centimeter. Reflection gratings are made by ruling the lines on a polished surface of speculum metal.

Let AB (Fig. 198) denote the enlarged cross section of a transparent grating, a, b, c, etc., the transparent lines. Suppose a parallel beam of monochromatic light of wave length λ incident normally on the grating. If the grating were

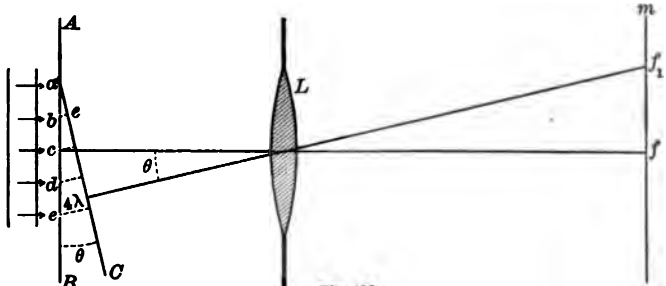


Fig. 198

absent, such a train of plane waves would be brought to a focus by the lens L at its principal focal point f. The interposition of the grating has no other effect on this image than to reduce its intensity. It does serve, however, to introduce other images which are not present before it is interposed. The explanation is as follows:

Each opening, a, b, etc., becomes a new source of secondary waves, and a surface may be drawn through a and so inclined to AB that it will touch all these secondary waves at points having the same phase of vibration. Thus the distance of b from aC is one wave length λ; of c, 2λ; of d, 3λ; etc.; and all the waves from the parallel apertures of the grating arrive at aC in the same phase. Hence aC is another plane wave which the lens will bring to a focus at f_1 on a secondary axis.

Similarly other lines may be drawn at such angles that be will equal 2λ, 3λ, etc.; and for each position the line will be a wave front for all the secondary waves arriving from the apertures of the grating, and this resultant wave will be brought to a focus in the focal plane of the lens.

For any inclination of aC making be greater or less than an exact multiple of λ, aC is not a wave front and the disturb-

ances from the openings of the grating arrive at the corresponding focus out of phase and suffer interference. For example, suppose be to be less than λ by one per cent. Then the waves from the fifty-first opening arrive at aC half a wave length ahead of those from the second; those from the fifty-second, half a wave length ahead of those from the third; etc. Hence at the focus there is complete extinction of the light by interference. Between f and f_1, therefore, there is a black band; and between the successive colored images, called images of the first order, second order, and so on, are dark spaces.

A series of nearly equidistant images of a distant source of light may then be obtained by means of a grating and a lens. While such images are most perfectly produced by precise artificial means, still they are by no means lacking in nature. The colors of changeable silk, of the feathers of some birds, and of mother-of-pearl are imperfect replicas of those given by a reflection grating. The colored rays which may be seen when the sun near the horizon shines through the foliage of trees a mile away are diffraction fringes. When a small distant source of light, an open electric arc for example, is viewed through a semitransparent screen of regular structure, such as a linen handkerchief or a silk umbrella held close to the eye, two series of images at right angles to each other may be seen. There are two series because there are two sets of cross threads in the fabric. In fact the eyebrows or fine feathers may serve as imperfect gratings to yield distinct images of a bright light at a distance.

306. Measurement of Wave Lengths by a Grating.—If θ_1 is the angle between the grating and the new wave front aC, which forms the image of the first order, that is, the angle between the direction of the incident light and that of the first image (Fig. 198), and if d is the distance ab from center to center of the openings in the grating, then the triangle abe gives the relation

$$be = \lambda = d \sin \theta_1. \qquad (68)$$

For the second image $$2\lambda = d \sin \theta_2.$$

From equation (68) it is obvious that the longest waves are found at the greatest deviations. For the first image, for which θ_1 is small, the wave lengths are nearly proportional to the deviations.

The procedure for measuring wave lengths by a *spectrometer* fitted with a plane grating may be described by reference to Fig. 199. At the

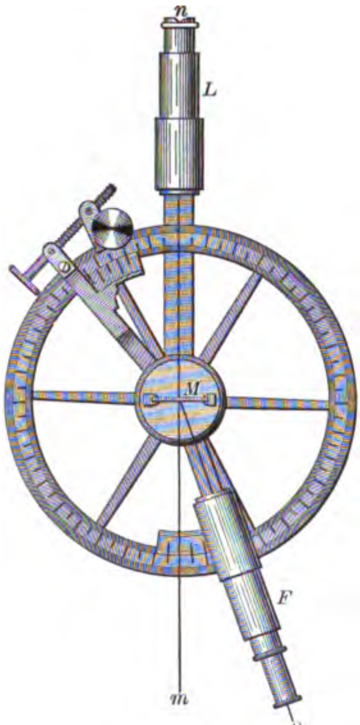


Fig. 199

outer end of the collimating telescope L is a narrow adjustable slit n, which must be strongly illuminated by the D light of burning sodium, or the red light of burning lithium, for example. The lens at the other end of this telescope converts the divergent waves from the slit into the plane waves of a parallel beam. This beam is incident normally on the grating at M, the lines of which are parallel to the slit.

The view telescope F is first set so that its cross wire coincides with the first colored image of the slit given by diffraction to the right of the line nm, and the reading on the circular scale is noted. The telescope is then shifted until its cross wire coincides with the first image to the left of nm, and the reading is again taken. Half the angle between the two sighted positions of the view telescope, equal to half the difference between the two readings on the circular scale, is the angle θ_1. The grating space d may be obtained from the maker of the grating, or it may be measured directly by a micrometer microscope.

To illustrate by an example.

	Reading to the right	72° 41′
	Reading to the left	55° 28′
Therefore	$\theta_1 = 8° \; 36.5'$.	

The distance d for the grating was 0.000446 cm. Substituting in formula (68), $\lambda = 0.000446 \times 0.149679 = 0.0000667$ cm.

307. Dispersion by a Grating. — When a grating is illuminated by white light instead of monochromatic light, the succession of images of one color gives place to several colored bands on both sides of the central white image of the source. The white light undergoes dispersion, the colors in each

band appearing in the order of their wave lengths, the violet nearest the central image and the red farthest away. Such a spectrum produced by a grating is called a *normal spectrum;* the dispersion is due to diffraction and interference.

It will be recalled that prismatic dispersion is due to the different velocities with which the several colors are transmitted through transparent media (§ 294). Dispersion by a grating is due to diffraction and is dependent entirely on wave length. It is obvious from equation (68) that there is a different value of the deviation θ for each different wave length. Now since white light is composed of all the spectral colors, or the vibrations of all wave lengths which excite vision, it follows that when white light is incident on a grating, it will form a succession of colored images, one for each wave length, and these will shade off from one to the next in the order of the spectral colors. Violet is the psychological response to the shortest waves which excite the sensation of color, and red to the longest waves.

The spectrum produced by a grating is said to be *normal* because, when θ in the first spectral band is small it is proportional to λ (68), or the angular deviation of each color from the direction of the incident light is directly proportional to its wave length.

A pure spectrum is one in which the colors do not overlap. The spectrum of the first order is the only pure grating spectrum. The wave length of extreme red is about 0.00076 mm.; of extreme violet, 0.00039 mm. Then

For the red of the 2d order $\qquad d \sin \theta = 2 \times 0.00076$

For the violet of the 3d order $\qquad d \sin \theta' = 3 \times 0.00039$

Therefore $\qquad d \sin \theta > d \sin \theta',$

or the red of the second order overlaps the violet of the third. Grating spectra are limited to the first three orders, the overlapping in higher orders being sufficient to reproduce white light.

The mean of the wave lengths for extreme red and extreme violet is 0.00054 mm. This is the wave length of the yellow between the Fraunhofer lines D and E. Yellow is therefore at the middle of all grating spectra, while in prismatic spectra it lies much nearer the red end. Figure 200 represents two

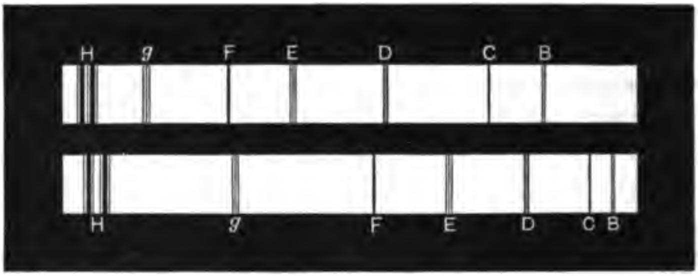


Fig. 200

spectra of equal length, the upper one a grating spectrum and the lower a spectrum given by a flint glass prism. It will be noted that the D of the one corresponds very nearly with the F of the other; also, that in the grating spectrum the red end is relatively more extended than in the prismatic spectrum, while the violet end is less.

308. Wave Lengths and Frequencies of Vibration. — The unit commonly employed in measuring wave lengths of light is the *tenth-meter*, of which 10^{10} are required to make a meter. The following are the values for the principal Fraunhofer lines in air at 20° C. and 760 mm. pressure:

A	7621.31	E_1	5270.52
B	6870.18	E_2	5269.84
C	6563.07	F	4861.51
D_1	5896.18	G	4340.63
D_2	5890.22	H_1	3968.62

Taking the velocity of light as 300 million meters a second, or 300×10^{16} tenth-meters, the frequencies of vibration corresponding to the above spectral lines may be found by dividing this velocity by the several wave lengths, since $n = V/\lambda$. The following are the results:

A	393.6×10^{12}	E_1	569.2×10^{12}
B	436.7×10^{12}	E_2	569.8×10^{12}
C	457.1×10^{12}	F	617.1×10^{12}
D_1	508.8×10^{12}	G	691.1×10^{12}
D_2	509.3×10^{12}	H_1	756.0×10^{12}

Thus the light entering the eye and producing the color of violet represented by the line H_1 is due to 756 millions of millions of vibrations a second. A photograph of the sun has been taken with an exposure of only one twenty-thousandth of a second. During this short period a beam of light 15,000 meters (9.32 miles) in length enters the camera, and fully 375×10^8 or 37,500 millions of waves of violet light make their impression on the sensitized plate.

II. EMISSION AND ABSORPTION OF RADIATION

309. Types of Spectra. — The methods of analyzing radiation by the dispersion produced either by a prism or a diffraction grating have already been described. The spectra from different sources obtained by these methods fall into two general classes :

A. Continuous Spectra. The radiation from a hot solid or liquid forms a *continuous spectrum* without any interruptions in the succession of colors and wave lengths from one extreme of visibility to the other. This is true however narrow the slit and pure the spectrum. Light emitted by a white-hot solid or liquid is composed of radiations of all possible wave lengths between red and extreme violet. Such spectra are not characteristic of the substances producing them, but only of their temperatures.

A continuous spectrum is given by the white-hot positive carbon of an electric arc lamp, by all gas and candle flames, and by molten metals. The light emitted by flames of burning carbon compounds comes from the minutely divided solid carbon set free by the combustion of the gas or vapor and raised to incandescence by the intense heat of the flame.

All bodies when heated begin to glow at about the same temperature in the neighborhood of 400°. They become visible in the dark at first as a dull red. As the temperature

rises other colors are added to their spectrum in the order of wave lengths, violet appearing only at the highest temperature. Conversely, when the continuous spectrum of the radiation from the positive carbon of an arc light is projected on a white screen, and the current is then cut off, as the carbon cools, the violet, blue, green, yellow, and red disappear in succession.

It has already been pointed out that not all self-luminous bodies are hot. Attention has been directed to the fact that there are many cases of faintly self-luminous bodies at low temperatures (§ 254). Certain glowworms, fireflies, bacteria, and beetles emit a faint light, the wave lengths of which in general all fall within the limits of the visible spectrum. The luminous efficiency of this light is therefore very high as compared with artificial sources of illumination.

B. Discontinuous Spectra. The spectra of incandescent gases and vapors consist of a limited number of separate bright lines, each of which is an image in one color of the slit. These are called *bright-line spectra* to distinguish them from the discontinuous spectra crossed with dark Fraunhofer lines, such as those of the sun and fixed stars.

Bright-line spectra are produced by volatilizing metallic salts in a Bunsen flame, by sending electric discharges through rarefied gases in glass tubes, or between metallic electrodes in the air. The characteristic spectrum of the luminous vapor of sodium is the yellow line *D*. With a spectroscope of sufficient resolving power this one line is resolved into two very near together, and each of these is found to be double. No two vapors give rise to the same series of bright-line images. The number varies greatly, ranging from about ten in the case of the alkali metals to several thousand in the spectra of iron and uranium.

The lower spectrum in Figure 201 (from the Lick Observatory Bulletin, No. 62) is the bright-line spectrum of hydrogen in the G-H region; the upper one is the bright-line spectrum of iron vapor in the same region.

The fact that all gases and vapors give discontinuous spectra was definitely established by Bunsen and Kirchhoff in 1860. By means of this differentiation between the spectra

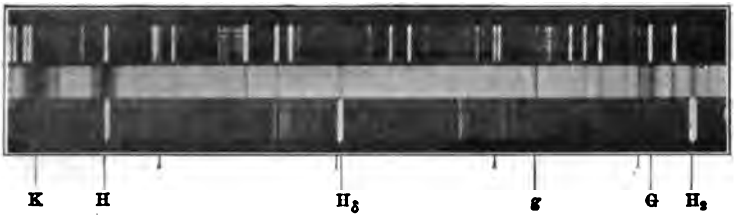


Fig. 201

of different metallic vapors, they discovered the rare metals rubidium and cæsium. The same method of spectrum analysis has been applied in recent times to the discovery of helium and some rare atmospheric gases.

Bright spectral lines differ widely in width and intensity. This may be true even of lines due to the same element. Wide lines are not really monochromatic, nor are even the narrow ones strictly so. It has been shown by means of instruments of extreme resolving power that some apparently narrow lines are composed of a number of still narrower components, which ordinarily overlap.

310. Absorption Spectra. — A pure solar spectrum, crossed by the dark Fraunhofer lines, is a discontinuous spectrum, but it belongs to a subclass known as *dark-line* or *absorption spectra*. These are made discontinuous by losses due to absorption of radiation in the passage through transparent media. The absorption producing the dark lines of the solar spectrum takes place chiefly in the outer envelope of the sun's atmosphere. The incandescent photosphere of the sun, which would by itself present a continuous spectrum, is surrounded by a mass of gases and vapors through which the radiation from the photosphere must pass. Absorption takes place in this reversing layer; and the dark lines, which are only relatively dark in comparison with the adjacent brighter

portions of the spectrum, are the inversion of those luminous radiations which form the emission spectra of the gases and vapors in the sun's outer envelope.

The principle of absorption is the same as that of resonance or co-vibration in sound. Every gas or vapor when white-hot emits radiations of the same wave length as those which it absorbs from an independent source when at a lower temperature. Thus the *D* lines of the solar spectrum coincide exactly with the bright lines given by sodium vapor in a state of incandescence. Not only has the *coincidence* been established between the Fraunhofer *D* lines and the yellow lines of luminous sodium vapor, but the *reversal* of these lines by sodium vapor as the absorbing agent has been accomplished. These results laid the foundation for the science of spectrum analysis, by which the approximate composition of self-luminous celestial bodies has been determined.

The presence of hydrogen and iron in the sun's atmosphere is demonstrated by the correspondence between the bright lines composing their spectra and dark absorption lines in the middle spectrum of Figure 201, which is that of the sun.

Kirchhoff established in 1860 the following law of spectrum analysis when the radiation is a pure temperature effect:

The relation between the emissive power and the absorbing power, for any definite color or wave length, is the same for all bodies at the same temperature. If light from a luminous vapor at a higher temperature traverses the same vapor at a lower temperature, the light absorbed in the latter is greater than the light emitted by it. The result is relatively dark lines, or a reversal.

The reversal of the sodium line may readily be shown by means of a direct current arc lamp and a hollow prism filled with carbon disulphide. The positive carbon should be larger than the negative and should be placed below. In a cup-shaped cavity in its upper end is placed a small piece of metallic sodium. With a rather wide slit in front of the lantern condenser, the light of the burning sodium when the

current is turned on gives a broad yellow band on the screen. When the sodium vapor in the arc is most copious, the middle of the yellow band turns dark on account of absorption by the outer and cooler layer of sodium vapor surrounding the arc. This phenomenon is known as self-reversal by absorption. It is especially pronounced in

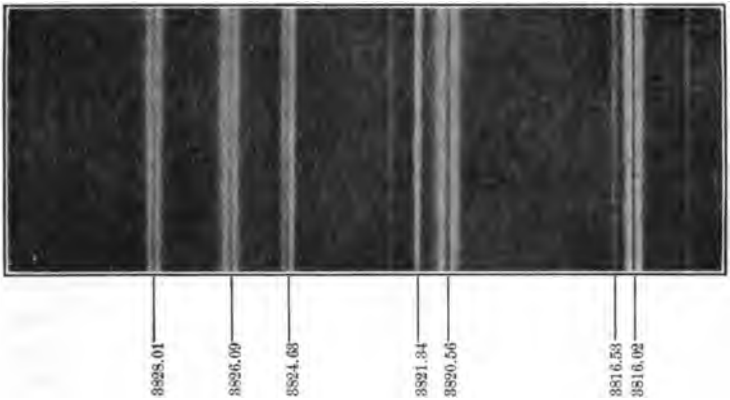


Fig. 202

spectra given by the electric arc. Figure 202 shows the self-reversal in a portion of the arc spectrum of iron given by a grating. The fine dark lines in the middle of the bright bands are due to self-reversal.

311. Infra-red and Ultra-violet Radiation. — The radiation affecting the eye and forming the visible spectrum lies between wave lengths of about 7500 and 3900 tenth-meters. The corresponding range of vibration frequencies is therefore a little less than would be called an octave in sound.

Besides these waves which give rise to sensations of color through the eye, there are many others which are entirely invisible, and yet possess all the physical characters of light waves. They may be reflected, refracted, diffracted, and polarized the same as waves exciting vision. Physically they differ only in wave length from those giving rise to optical

perception. The eye with its limited range of sensibility is not directly affected by them; but when other tests for detecting and registering waves are applied, there is no line of demarcation between the visible and the invisible radiations. The invisible portion of the solar spectrum is crossed by absorption lines precisely as in the visible part. (See Fig. 188.)

The radiation beyond the red end of the visible spectrum and of wave length greater than about 7500 is known as the *infra-red;* that beyond the extreme violet and of wave length shorter than about 3900, as *ultra-violet.* Both the infra-red and the ultra-violet radiations must be investigated by other physical effects than sight.

To explore the infra-red end of the spectrum resort is had to the heating effect of radiation. Explorations by means of the thermopile (§ 573) and by Professor Langley's bolometer show that the region of maximum heating effect is a little way into the infra-red region, and at wave lengths about 9500 to 10,000 tenth-meters. Langley measured lunar radiation having a wave length of 170,000 tenth-meters (0.017 mm.), or more than twenty times as long as the longest waves exciting vision. Professor Rubens has measured infra-red waves as long as 240,000 tenth-meters (0.024 mm.). These waves were thirty-two times as long as the longest ones affecting the eye as extreme red. They extend five optical octaves below the visible spectrum.

Ultra-violet radiation is especially effective in producing photographic action, particularly with silver salts. Photography is therefore the most satisfactory method of exploring the ultra-violet end of the spectrum. It cannot be employed for the infra-red because it is difficult to make a film that is sensitive as far down even as the visible red extends. Rowland succeeded in photographing from wave length 7000 tenth-meters in the red to 3000 in the ultra-violet. Ultra-violet radiation has been photographed to wave length 2020, or to a frequency about twice as great as that of extreme violet light. By other means it has been detected to wave

length 1000 tenth-meters, or to two optical octaves above the visible spectrum.

Figure 203 represents the entire spectrum from extreme infra-red up to wave length 3000 in the ultra-violet, the limit of Rowland's photographs. It is drawn to an even scale of frequencies instead of wave lengths. The extreme ultra-violet detected at wave length 1000 tenth-meters, or 100 micro-millimeters, with a frequency of 3000 billions of vibrations a second, lies twice the length of Figure 203 beyond its short wave length end. A scale of even wave lengths would make a spectrum more than sixty times the length of the visible spectrum.

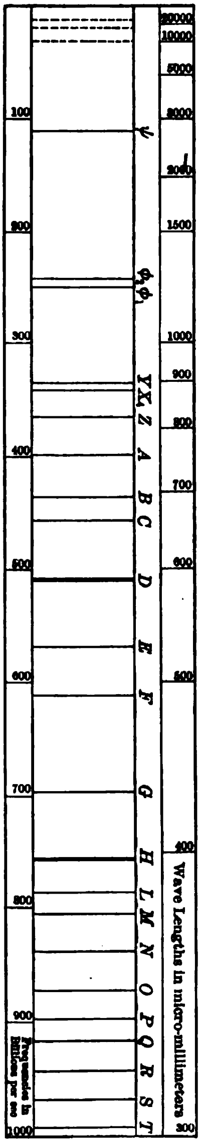


Fig. 203

Only a few of the principal absorption lines are represented in Figure 203. The last two dotted lines in the infra-red mark the limit reached by Langley and Rubens, respectively.

312. Reëmission of Absorbed Radiation. — In general it may be said that a part of the radiation absorbed by a body is again emitted, but always with a change of wave length. The reëmitted radiation is not due to high temperature and does not follow Kirchhoff's law. The reëmission as waves too long to affect the retina by a black body absorbing light and transforming it into heat is a familiar case in point. In many cases the transformed radiation falls within the range of visual perception. The phenomenon is then called *luminescence*, a name applied by Wiedemann to cover all cases of visible radiation not directly due to high temperature.

Wiedemann recognized eight causes of luminescence, including radiation, chemi-

cal action, heat, friction, and X-rays. Attention is directed here chiefly to the first of these.

Many substances may be stimulated into giving out light by exposing them to ultra-violet radiation. In all such cases the wave length emitted is longer than that of the stimulating radiation. Photo-luminescence, or luminescence stimulated by light, is subdivided into two groups, *fluorescence* and *phosphorescence*, the essential difference between the two depending upon the duration of the luminescence after the stimulus is withdrawn.

313. Fluorescence.—The emission of light by a body stimulated by radiation of a different wave length, in so far as the emission continues only during the period of stimulation, is called *fluorescence*. Fluorspar was the first substance in which the phenomenon was noticed; hence the name fluorescence.

Many substances exhibit this property. Uranium glass, solutions of the sulphate of quinine, eosin (a product of coal-tar distillation), and petroleum possess the property of fluorescing in varying degrees. When a beam of sunlight passes through a green solution of chlorophyll, its path is marked by a bright red streak. Uranium glass is yellow, but it has a beautiful green surface tint. Petroleum when strongly illuminated and viewed obliquely shows a splendid blue color.

Most of these substances have the property of absorbing ultra-violet radiation and transforming it into longer waves visible to the eye. When a beam of light from an electric lantern is passed through a dark violet glass, it loses nearly all visible light; but if it is then directed upon a block of uranium glass, the invisible waves are transmuted into a brilliant green, and the glass stands out vividly in the darkness. A vessel of kerosene oil in the ultra-violet beam appears azure blue. These substances shine most brilliantly in ultra-violet light just beyond the violet end of the visible spectrum.

Quinine in solution, on the other hand, possesses the property of converting long infra-red waves into shorter ones visible to the eye. This and a few other aqueous solutions are exceptions to the general law that reëmission increases the wave length of the radiation.

It is possible by means of fluorescence to photograph invisible objects. Thus, an inscription written on white drawing paper with colorless sulphate of quinine dissolved in a solution of citric acid is invisible in white light. This substance fluoresces and so transmutes the highly active ultra-violet radiation into longer waves, which do not act so vigorously on a photographic plate. When therefore such a sheet is illuminated by means of the arc lamp and is photographed, the parts to which the sulphate of quinine has been applied come out darker than the untouched surface of the paper (Fig. 204). With one or two sheets of blue glass in front of the lantern, the inscription written with the sulphate of quinine solution stands out in nearly white letters on a blue ground.

Fig. 204

314. Phosphorescence.—In fluorescence the emission ceases as soon as the stimulating radiation ceases to fall on the substance. But many substances, notably the sulphides of calcium, of barium, and of strontium, continue to emit light for appreciable periods after stimulation is withdrawn. Bodies which thus store radiant energy and continue to give it out as light afterwards are said to be *phosphorescent*. The name is derived from the similar appearance of phosphorus during slow oxidation. Some diamonds have this property, but sulphide of calcium, especially when mixed with small quan-

tities of other substances, exhibits this property in a most extraordinary degree. Balmain's paint is a preparation mainly of sulphide of calcium and a trace of bismuth. The excitation due to sunlight during the day enables it to give out light all night without losing its whole store of luminous energy. Even after having been kept in darkness for many months, a sheet of luminous paint may give off enough invisible radiation to fog a photographic plate.

The warming of a sheet of luminous paint causes it to shine more brightly. Conversely, chilling it dims its light. On the other hand Professor Dewar has discovered that many substances acquire the property of phosphorescence only when cooled in liquid air to a temperature of about 200° C. below zero. Such are gelatine, horn, paper, ivory, and egg shell. At this low temperature they absorb radiant energy and store it for emission as light when warmed.

A curious instance of phosphorescence is afforded by the termite or white ant hills in the region of the Amazon. These hills are from five to ten feet high, are made of hard clay, and are bare of vegetation. In the night they glow and scintillate with a shifting phosphorescent light, which is visible at a distance when the night is dark. The weight of opinion appears to be that the phosphorescence is in the hills themselves and not in the insects.

315. Applications of Doppler's Principle. — The Doppler effect in light (§ 214) is the apparent change in wave length produced by the relative motion of the observer and the source in the line of sight. When the observer and the source of light are approaching each other, the apparent wave length is shortened, and *vice versa*. This apparent shortening or lengthening of the waves gives rise in discontinuous spectra to the shifting of the lines toward the violet end or the red end of the spectrum, respectively. Since the velocity of light is so very great, a large velocity of approach or recession is necessary in order to produce a measurable shift in the position of spectral lines.

Referring to § 214, the apparent wave length is

$$\lambda' = \frac{V \mp v}{n} = \lambda \mp \frac{v}{n}.$$

Whence $$\pm v = n(\lambda - \lambda') = \frac{V}{\lambda} \Delta\lambda,$$

where $\Delta\lambda$ is the apparent change in wave length, and v the relative velocity of approach or recession. It is easy to calculate what this velocity must be to produce any assumed change in apparent wave length, 0.1 of a unit in tenth-meters, for example. Take V, the velocity of light, as 300,000 km. a second, and λ as the wave length of the F line of hydrogen, 4861 tenth-meters. Then

$$v = \frac{300000000}{4861 \times 10^{-10}} \times 0.1 \times 10^{-10} = 6170 \text{ m.} = 6 + \text{km.}$$

per second.

The linear velocity of rotation of a point on the sun's equator is 2 km. a second. The difference in velocity of points at the two ends of the solar equator with respect to an observer is then 4 km. a second. Vogel, in 1871, succeeded in detecting the displacement of Fraunhofer lines by comparing the spectra from the two limbs of the sun, and Young determined the sun's period of rotation by the same method.

Some of the dark lines in the spectra of sun spots are not infrequently broken and greatly distorted. This distortion is due to the swift motion of matter, especially hydrogen, in the line of sight. The broken lines of Figure 205 are three views of a hydrogen line taken at intervals of about four minutes. In this particular case the gas was projected outward toward the observer with a velocity of about 480 km. (300 miles) a second.

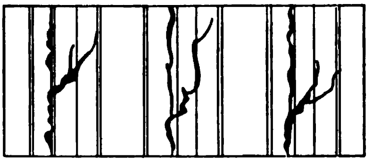


Fig. 205

By the refined measurement of the displacement of Fraunhofer lines in the photographs of stellar spectra, the velocities of stars in the line of sight may be determined with an error not exceeding one tenth km. per second. The investigation of a great many stars shows that most of those in one direction in the heavens have an apparent motion toward the sun, and those in the opposite direction an apparent motion away from the sun. The inference is that the solar system is moving through space in the other direction.

316. Spectroscopic Binary Stars.— One of the most interesting applications of the Doppler principle is the investigation of very close double stars, which are entirely unresolvable by the most powerful telescopes. These are known as spectroscopic binary stars.

When light comes from a double star, both members of which are self-luminous, the Fraunhofer lines in the stellar spectrum become double, separate to a maximum distance, and then come together again. The complete cycle of these changes takes place in the period of revolution of the components of the double star about their common center of mass.

The lines in the spectrum of such double stars appear single when one member is behind the other in the line of sight; but when the two components of the binary system are side by side, one is approaching the observer and the other receding from him. Any given Fraunhofer line coming from the two sources is then shifted in opposite directions and appears double.

A more interesting type still is that of spectroscopic binaries in which one member is dark. These are usually variable stars, such as Algol. By means of the displacement of lines in its spectrum it was found that Algol varies in its velocity in the line of sight, positive and negative motions alternating with each other, the period of these motions agreeing with the period of the variable luminosity. It has

thus been shown that Algol is a binary system composed of one bright and one dark member, both rotating around their common center of mass.

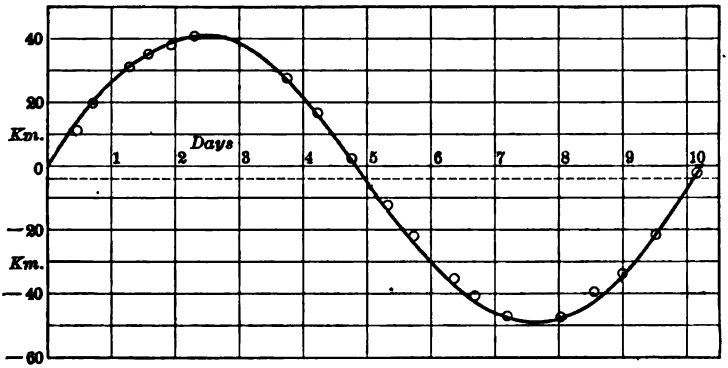


Fig. 206

Other stars of this type have been investigated. When the observations are plotted with times as abscissas and radial velocities (velocities in the line of sight) as ordinates, the resulting curve is a sine curve. Figure 206 was plotted

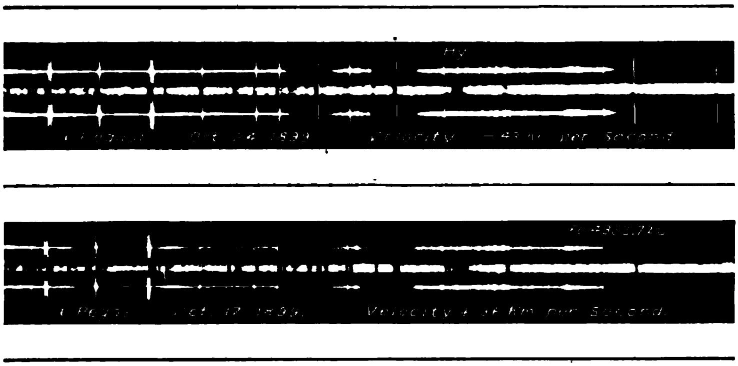


Fig. 207

from data derived from some of the observations on the star ι *Pegasi* made at the Lick Observatory. The curve shows

that the orbit of this star projected on a plane containing the line of sight is almost perfectly circular, and that the star is the luminous member of a binary, the other member of which is dark. The dotted line denotes the center of mass of the system. The period of revolution in the orbit is 10.213 days, and the diameter of the circular orbit is 13,480,000 km.

The displacement of stellar lines in the spectrum of this spectroscopic binary is shown in Figure 207. The middle spectrum is that of the star, and the comparison spectra are the bright lines given by titanium. In the upper group the star is receding from the observer, and in the lower one it is approaching. The stellar lines are displaced in opposite directions in the two groups.

Problems

1. If the grating space $d = 589 \times 10^{-5}$ mm., and $\theta_2 = 11^\circ 32'$ for D light, what is its wave length [equation (68)]?

2. If D light falls normally on the grating (Fig. 198), for which $d = 589 \times 10^{-5}$ mm., and the screen is 5 m. from the lens L, find the distance ff_1 for the first bright line image.

3. If the angle θ_2 for the blue line of cadmium ($\lambda = 48 \times 10^{-5}$ mm.) is 12°, calculate the grating space d (68).

4. The velocities of the source in the line of sight were 36 and -43 km. per second respectively for the two spectra of Figure 207. What were the corresponding displacements of the line $\lambda = 4383.7 \times 10^{-10}$ m., in terms of the unit of wave length, 10^{-10} m.?

5. The maximum computed velocities of ι *Pegasi* in the line of sight were 43.7 and -52.1 km. per second (§ 316). Calculate the diameter of the orbit of the visible member of this binary on the assumption that its orbit is a circle with the line of sight in its plane.

CHAPTER X

SENSATIONS OF COLOR

317. Modes of Producing Color.—Color itself has no objective existence. It is the response of sensation to the stimulus of light. The only physical difference corresponding to different simple colors is a difference of wave length. The extreme red of the spectrum is a sensation excited by the longest ether waves affecting the eye; extreme violet, by the shortest. The production of spectral colors from white light involves some process of separating the waves of different length, so as to get their individual effect. Such a separation we know is brought about by dispersion and interference. Other modes of producing colors from white light are reflection and selective absorption. Then there are other sensations of color depending on over-stimulus or fatigue of the retina.

Although each light wave of different length produces a definite simple color sensation, it is not true that every mixed color sensation is produced only by one set of waves of definite frequencies. The combined effect of all the visible light waves from a white-hot solid is the sensation of white; but it is also possible to produce the sensation of white light by combining only two kinds of waves. For example, the superposition of light from the red end of the spectrum and that from a region between the green and the blue also produces the sensation of white. A similar sensation is excited by the combined effect of violet and yellowish green. Any two colors, the combined effect of which is the sensation of white, are said to be *complementary*.

318. Selective Absorption. — Substances which absorb some radiations and transmit or reflect others are said to exercise *selective absorption.* The absorption of radiation is usually selective. Familiar examples occur within the limits of the visible spectrum. Thus, red glass colored with the suboxide of copper transmits red and absorbs all other visible radiation. For this reason "ruby glass" is used in photographic rooms; it does not transmit the radiations which fog a sensitized plate.

Many colored liquids exhibit the same peculiarity of selective absorption. Thus amyl alcohol, dyed with aniline red, cuts off by absorption every tint except red; cupric chloride dissolved in hydrochloric acid cuts off all except green and some blue.

Many substances are quite transparent within the limits of the visible spectrum, but show selective absorption in the infra-red and ultra-violet regions. Glass, for example, is opaque to waves shorter than 3500 tenth-meters, and longer than about 30,000 tenth-meters. Quartz is transparent to radiations between wave lengths 1800 and 70,000; rock salt between 1800 and 180,000; while fluorite transmits the shortest known ultra-violet waves, and down to wave length 95,000 in the infra-red.

A body which shows selective absorption in the infra-red or ultra-violet only appears colorless, but in a physical sense it is similar to a colored body. If we obtained visual sensations from ultra-violet radiations, glass, which absorbs ultra-violet strongly, would not appear transparent and colorless. The insensibility of the eye to ultra-violet light appears to be due to the fact that these radiations are absorbed by the media of the eye before they reach the retina.

319. Color of Opaque Bodies. — All bodies, except those with highly polished surfaces, reflect light by irregular reflection, and in most cases the light penetrates more or less into the medium before it is diffused. If all the constituents of white light are reflected in the same proportion, the body appears white or gray. Such is the case with a sheet of white paper or a white screen on which the solar spectrum is projected. White bodies reflect diffused light in all directions, and without preference for light of particular wave

lengths. Hence all the colors of the spectrum on such a screen appear the same as when they are received directly into the eye placed in the path of the beam diverging from the prism.

If, however, a body exhibits selective absorption for some frequencies lying within the visible spectrum, then the light reflected from it consists of a mixture of the color components of white light which the body reflects. The colors of opaque bodies are therefore chiefly the residuals left after absorption.

The illumination of opaque bodies by colored light is very instructive. Project a spectrum of the sun or of the electric arc, with a rather wide slit, on a white screen in a dark room, by means of a carbon disulphide prism. Select a flower with rich red petals in large masses, such as the tulip or certain geraniums, and pass it through the different colors of the spectrum on the screen. In the red the flower will shine with its usual bright red color; but as it is passed along into the green, it becomes black, and shows no power of reflection for the remaining colors of the spectrum. All the colors except red are almost completely absorbed. The red, on the contrary, is reflected, and gives color to the body. A piece of ordinary red flannel is brilliantly red in the less refrangible end of the spectrum, but suddenly turns to a dirty brown, and then a dead black, when moved away from the red toward the violet.

It is obvious that the color of a body does not inhere in the body itself, for it can exhibit no color not already present in the light which illuminates it. With homogeneous illumination differences of color are no longer possible. This fact is strikingly illustrated by viewing objects of various colors in a room lighted only with burning sodium. The most healthful face is of an ashen hue, and brilliant flowers are reduced to a faded yellow. It requires the white light of the sun, in which innumerable colors are blended, to disclose to our eyes the variegated tints of nature.

320. Color of Transparent Bodies. — A body transparent to certain radiations affecting the eye, and not to others, appears colored by transmitted light, and the color is due to the mixed impression excited by the transmitted radiations. The colors of transparent bodies are due to their power of selective absorption. If a piece of blue cobalt glass be interposed in the solar beam, the spectrum will consist of a small

amount of extreme red and all of the indigo-violet (Fig. 208, 2). A piece of glass, colored red with the suboxide of copper, allows only the red and orange-red rays as far as the *D* line to pass through (Fig. 208, 1). All the rest of the spectrum is completely stopped by this glass. If now the light be passed in succession through the copper red and the cobalt blue glasses, the only spectral color surviving the process of double absorption will be the extreme dark red below the line *B*, which is transmitted by both.

A solution of potassium bichromate transmits the less refrangible part of the spectrum only up to the Fraunhofer line *b*.

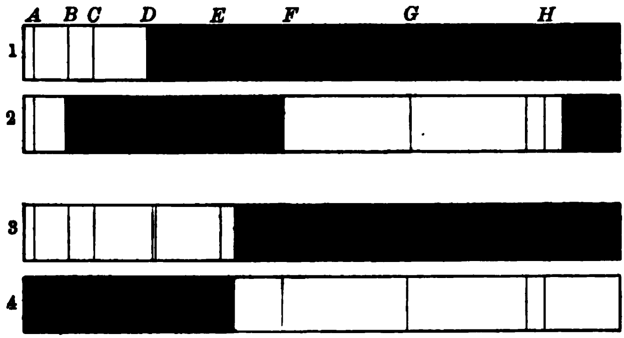


Fig. 208

A solution of the ammoniated oxide of copper transmits only the more refrangible part of the spectrum from the *b* line on (Fig. 208, 3 and 4). These two colors, therefore, contain all the spectral tints, and are complementary to each other. But if the two solutions in flat glass cells be placed in the path of a beam of sunlight, one behind the other, the combination scarcely permits the passage of any light at all. The light that struggles through the one solution is stopped by the other.

If the blue ammoniated copper oxide solution be placed in front of a yellow solution of normal potassium chromate of the proper density, the light transmitted by the two will be

green. If their separate spectra be examined, it will be found that both solutions transmit green. Green is the only color common to the two, and the only one not stopped by absorption in the one solution or the other.

321. Surface Color.—Some thin metallic films and solid aniline dyes reflect one color and transmit another. Thus gold reflects characteristic yellow light, but a very thin sheet of gold leaf is green by transmitted light. The reflected and transmitted colors are complementary.

Fuchsine has a superficial metallic green color; an alcoholic solution of it forms a brilliant red dye and transmits no green. In such cases the surface color appears to be due to selective reflection at the surface, the other colors being transmitted through a thin layer or absorbed by a thick one.

Silver exhibits similar properties if we include the ultra-violet. The surface color of silver is gray. A thin film of silver deposited on glass transmits nearly white light, and the ultra-violet is transmitted through a film so thick that no visible light gets through. Hence, with a quartz lens heavily silvered so as to be opaque to spectral colors, photographs may be taken by means of the transmitted ultra-violet rays. In such a photograph any surface which absorbs ultra-violet light, such as a surface covered with a fluorescent substance, appears dark or nearly black. A polished silver object shows dark because it absorbs the ultraviolet rays by which the photograph is taken.

322. Mixing Colored Lights.—For the perception of the mixed effect of two or more colored lights, it is necessary that they reach the retina either simultaneously or in quick succession. Visual impressions persist for a small fraction of a second; and, if one remains until after the arrival of another, both impressions are present at the same time.

If two partially overlapping disks of light be projected on a screen, and transparent colored bodies be placed in the path of the two beams, the light reflected to the eye from the overlapping area will consist of a real mixture of the two colored lights. Thus, if the ammoniated copper oxide solution be placed in the path of one beam and the potassium chromate in the other, the area common to the two disks will be white or gray, with the proper density of the two solutions. When

these colored lights are *added*, they cannot in any way be made to produce green. So the cobalt-blue and the copper-red glasses will give beams of light which by addition produce white.

Fig. 209

If a disk of cardboard, colored in sections, be rapidly rotated (Fig. 209), the result is a mixture or a superposition of visual impressions. But very different mixtures may produce the same visual impression. The eye has no power of analyzing light into its constituent colors. It can tell nothing about the composition of colored light. Colored light must be analyzed by means of a spectroscope armed with a prism or a diffraction grating.

323. Pigment Colors. — The effect of mixing pigments is not at all the same as that of mixing colored lights. In the case of pigments, the light reaching the eye is white light which has lost some of its components by absorption within the pigment which it slightly penetrates. A pigment color is the residual left after this absorption, and the mixture of two pigments gives rise to a double absorption instead of the addition of two colored lights.

The result of mixing pigments is the same as the passage of the light in succession through two colored glasses or colored liquids. Blue paint absorbs nearly every color except the blue and some green; yellow paint absorbs all but the red, yellow, and green. Therefore with white light incident on a mixture of fine yellow and blue pigments, the only color that escapes absorption is the green. If the ammoniated copper oxide solution and the potassium chromate solution be mixed, the transmitted light will be green, the same as when the light is transmitted through the two solutions in succession.

324. The Three Primary Color Sensations. — In the theory of color vision originated by Thomas Young and later elaborated

by Helmholtz, all color sensations are referred to three simple primary sensations of color, which have their counterpart in the nerve terminals of the retina. These are the sensation of *red*, the sensation of yellowish *green*, and the sensation of *blue-violet*. Red light stimulates but one of these sensations in the nerves of the eye. Yellow light stimulates two, the red and the green. The recognition of different colors is due to the excitation of the three primary sensations in varying degrees.

With three colored circles of red, green, and blue-violet light partially overlapping on the ~~screen (Fig.~~ 210), the resulting mixtures are readily made out. The overlapping of the red and green gives yellow; of the green and the blue-violet, peacock blue; of the blue-violet and the red, purple; while the overlapping of the three colored lights gives white. The tint produced by the overlapping of two of the primary color sensations is complementary to the third.

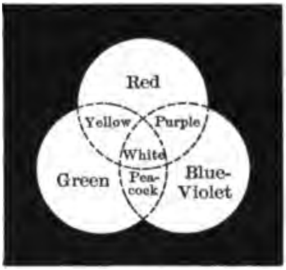


Fig. 210

The theory of the three primary sensations of color is the foundation of all the methods of color photography.

325. Color Blindness. — The Young-Helmholtz theory of color sensation affords a simple explanation of the facts of defective color vision known as *color blindness*.

A red-blind person is one who is entirely deficient in the nerve fibers by which the primary sensation of red is perceived. Only two primary color sensations remain for such an abnormal eye. In such cases red produces only a moderate stimulation of the nerve terminals giving green, and little or no stimulus of the blue-violet. Red therefore appears to such a person as a green of low luminosity. To such an eye there is no distinction between the green leaves of a cherry tree and the red ripe fruit. To a red-blind person the luminosity of the red end of the spectrum is much below normal.

Green blindness and blue blindness also occur, though rarely. To a green-blind person equal stimulation of the red and the blue-violet corresponds to the sensation produced by white. When such a color-blind

person looks at a spectrum of white light, he sees red at one end, violet at the other, and a white band between them.

By far the commonest form of defective color vision is red blindness. It occurs chiefly in the male sex, and affects three or four per cent of the community.

326. Subjective Colors. — The theory of primary color sensation furnishes a ready explanation of colors due to fatigue of the retina. It is well known that objects are quite invisible to one entering a faintly lighted room after exposure of the eyes for some time to bright light. After subjection to a strong stimulus the eye loses its sensitiveness to a weak one. When one enters from outdoors a room in which sensitized photographic films are cut and rolled, nothing is at first visible except a few dim red lights; but after a short time the eye recovers its sensibility; one then sees a room filled with machines and the operatives attending to all details of the work without difficulty.

This liability to fatigue is characteristic not only of the retina as a whole, but of any portion of it giving one of the primary color sensations. Fatigue of the retina causes it to lose the power of responding to the stimulus of any color long observed; and if the eye is then directed toward a dimly illuminated surface, a tint appears complementary to the one which has produced the fatigue. This is one form of subjective color. If the eye is fixed for half a minute on a colored picture, red for example, in a strong light, and is then directed toward a dimly lighted white wall, an image of the picture will be seen in green, enlarged if the wall is farther from the eye than the picture itself. The eye, fatigued for red, still retains its sensitiveness for green. Hence, the relatively faint white light is sufficient to stimulate for green, while it is an insufficient stimulus for red. The state of the retina serves to suppress the sensation due to a portion of the radiation, leaving the remainder.

It has been shown that when the eye is fatigued by white light it recovers its sensibility for different colors successively

after different intervals of time. If one looks out of a small window, such as a porthole, on a strongly illuminated sky, and then closes one's eyes, the after image of the window will appear in dissolving colors of brilliant hues.

Simultaneous color contrasts are another form of subjective colors. These are well displayed by laying thin tissue paper over black letters printed on green cardboard. The letters in a strong light are pink by contrast. The tissue paper furnishes a faint illumination compared to the strong green, and the unfatigued nerve terminals giving red cause the letters to appear pink in contrast with the complementary green. In this way certain colors are heightened by contrast, particularly if they are complementary.

CHAPTER XI

POLARIZED LIGHT

I. DOUBLE REFRACTION

327. Phenomena of Double Refraction. — In the study of single refraction in Chapter VIII it was tacitly assumed that when light enters a dense transparent medium it has but one velocity of transmission, and that there is consequently only one refracted ray for each incident ray. But all crystalline substances, except those whose fundamental form is the cube, possess the property of dividing each incident ray into two refracted ones. This phenomenon is called *double refraction.* It belongs also to many animal and vegetable substances and to homogeneous media, like glass, which are unequally strained in different directions. On account of this unequal strain the physical properties are not the same in all directions.

When a thin pencil of light traverses a double refracting body, it emerges as two pencils. One of these obeys the laws of single refraction (§ 277), and is called the *ordinary ray.* The other in general does not obey these laws, and is called the *extraordinary ray.*

In every double refracting crystal there is at least one direction in which light may pass without bifurcation. This direction is called the *optic axis.* The refracted rays diverge most widely when the incident pencil is perpendicular to the optic axis. Both rays then follow the ordinary laws of single refraction.

It should be noted that the angular separation of the ordinary and the extraordinary ray in many cases is so slight

that they are nearly superposed and not distinguishable as separate rays. The double refraction in such cases is clearly demonstrated by polarized light in a manner soon to be explained.

328. Double Refraction in Iceland Spar.—Iceland spar exhibits double refraction to a remarkable degree. It is crystallized carbonate of calcium, and occurs in rhombohedra, whose faces form angles with each other of 105° 5′ or 74° 55′. The angles of the crystal are the same in all specimens, but the lengths of the three edges may have any ratios whatever. Two of the solid angles at the extremities of the diagonal *ab* (Fig. 211) are contained between three obtuse plane angles. A line equally inclined to the three edges bounding one of these solid angles is the *axis* of the crystal. This axis is a definite *direction* rather than a definite line. Iceland spar has but one such axis, and hence is said to be *uniaxial*.

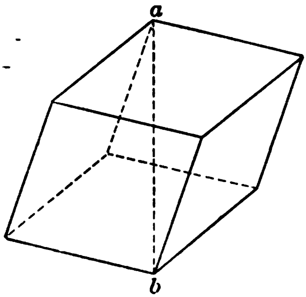


Fig. 211

Any plane containing the axis is a *principal plane.* If the two solid angles bounded by the three obtuse angles are cut away by planes perpendicular to the axis of the crystal, a parallel pencil of light, incident normally on either of these polished planes, will pass through without bifurcation.

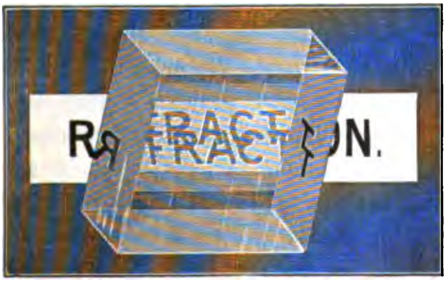


Fig. 212

If a crystal of Iceland spar be laid on a printed page, the letters will in general appear double (Fig. 212); or if a pencil of sunlight be admitted through a

small round opening in a shutter, and a crystal of Iceland spar be placed in the path of the pencil, so that it is normal to a face of the crystal, two equally illuminated disks will appear on the screen.

When the plane of incidence is perpendicular to the optic axis, both rays follow the laws of refraction; the index of refraction for the ordinary ray and the *D* line is 1.658; for the extraordinary ray it is 1.486. It follows that the extraordinary ray is transmitted through Iceland spar with greater velocity than the ordinary ray.

Crystals like Iceland spar, in which the index of refraction for the extraordinary ray is smaller than for the ordinary, are called *negative* uniaxial crystals. Quartz is also double refracting; but in quartz the angular separation of the two rays is smaller than in Iceland spar, the indices of refraction being 1.544 and 1.553 for the ordinary and the extraordinary rays respectively with sodium light. Such crystals, in which the index of refraction for the ordinary ray is the smaller of the two, are called *positive* uniaxial crystals.

329. Uniaxial Prisms.—A prism may be cut from a uniaxial substance either so that its optic axis makes equal angles with the faces which unite at the refracting edge, or so that it is parallel to the refracting edge. In the former case a pencil of light traversing the prism in the direction for minimum deviation will not divide into two rays, for its path through the prism is in the direction of the optic axis; but in the second case (Fig. 213) the plane of incidence is perpendicular to the optic axis, and the divergence of the ordinary and the extraordinary rays is a maximum. Two spectra will then be formed, the ordinary having the greater mean deviation for an Iceland spar prism, and the smaller for a quartz prism.

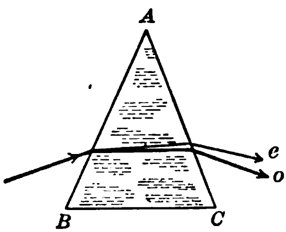


Fig. 213

Such a prism may be used to measure the refractive indices by setting it so that first the ordinary and then the extraordinary ray is at minimum deviation, and employing formula (57), § 282. The following are a few values for sodium light:

NEGATIVE CRYSTALS

	μ_o	μ_e
Iceland spar	1.658	1.486
Sodium nitrate	1.587	1.536

POSITIVE CRYSTALS

	μ_o	μ_e
Quartz	1.544	1.553
Ice	1.309	1.310

330. Wave Surfaces in Uniaxial Crystals. — If there is a radiant point within a transparent isotropic body, that is, one whose physical properties are the same in all directions, such as water or well-annealed glass, the wave surfaces about the radiant point are spheres, for the velocity of light is the same in all directions. If, however, the body is not isotropic, the velocity varies with the direction of transmission, and the wave surfaces are no longer spherical.

Now, in double refracting crystals there are two rays, and two indices of refraction; hence, there are two different wave surfaces; and, since the refractive index for the extraordinary ray is not the same in every direction, the velocity of this ray varies with the direction between limits. It follows that the wave surfaces for this ray are not spherical.

Huyghens demonstrated that in uniaxial crystals the wave surface for the extraordinary ray is a *spheroid* or *ellipsoid of revolution*, that is, the surface generated by rotating an ellipse about one of its axes coinciding with the optic axis of the crystal. If, further, there is only one ray in the direction of the optic axis, there is in general only one velocity in this direction and the sphere and spheroid touch each other on this axis.

There are, however, two cases corresponding to positive

and negative crystals. In the former the velocity of the ordinary ray is the greater of the two, and in the latter it is the less, except along the optic axis. In the former a section of the wave surfaces through the radiant point and containing the optic axis is like Figure 214 *a*; in the latter,

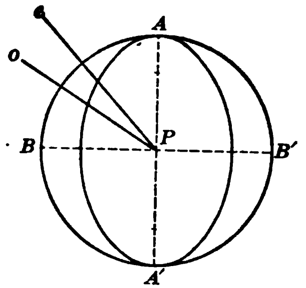


Fig. 214 *a*

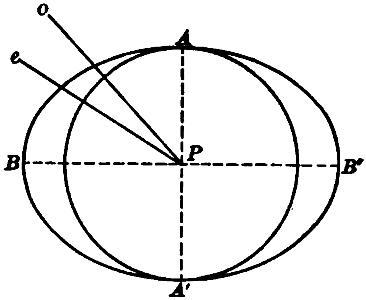


Fig. 214 *b*

it is like Figure 214 *b*. The two rays have the same velocity along AA', but unequal velocities in all other directions, the difference being the greatest along BB', perpendicular to the optic axis. Also, the direction of propagation of the extraordinary ray is not in general normal to its wave surface, while that of the ordinary ray is normal.

In certain uniaxial crystals, quartz for example, the velocity of the two rays is not the same even in the direction of the optic axis, the spectral lines in the spectroscope appearing double, indicating slight double refraction. In such crystals the two wave surfaces are not tangent to each other, but either the one lies wholly within the other, or the two intersect at four points symmetrically situated with respect to the optic axis.

331. Construction for the Refracted Rays in Iceland Spar. — The path of the two refracted rays in Iceland spar may be readily constructed from the two wave surfaces, especially in the selected cases in which both rays lie in the plane of incidence.

Let *ab* (Fig. 215) be the direction of the incident light and *bd* its plane wave front. *MN* is the surface of the spar. Also let the optic axis *bi* lie in the plane of the paper and therefore in the plane of incidence. The path of the ordinary refracted ray is found by the construction already employed in § 279. About *b* as a center, and with a radius which bears to *df* the ratio of 1 to 1.658 (v' to v), describe a circle; from *f* draw a tangent to this circle, and connect *b* and the point of tangency *o*. Then *bo* is the ordinary ray.

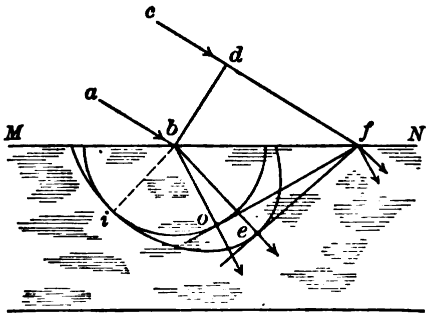


Fig. 215

Since the optic axis coincides with *bi*, the intersection of the plane of the paper with the wave surface of the extraordinary ray is the ellipse tangent to the circle at *i*. To find the direction of the extraordinary ray, draw from *f* the line *fe* tangent to the ellipse, and connect *b* and the point of tangency *e*. Then *be* is the extraordinary ray. The new wave fronts are *fo* and *fe*.

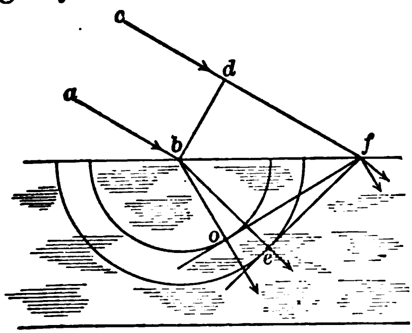


Fig. 216

As a second example, let the optic axis be perpendicular to the plane of the figure, and therefore to the plane of incidence. The intersections of the plane of the diagram with the two wave surfaces are then both circles (Fig. 216). From *f* draw tangents to both circles, and the lines *bo* and *be*, drawn from *b* to the points of tangency, are the directions of the ordinary and the extraordinary ray respectively. The new wave

fronts are *fo* and *fe*. This is the construction for the greatest divergence of the two refracted rays.

II. POLARIZATION BY REFLECTION

332. Plane Polarized Light.—A ray of common light has on all sides the same peculiarities. When a pencil of common light is received on a plane mirror, it is reflected, in whatever direction the plane of incidence may pass through the pencil. But there is a peculiar modification of light which has not the same properties on all sides; this asymmetry about the line of propagation is known as *polarization*, and light having this modification of asymmetry is called *polarized light*.

The term polarization was introduced by Malus in 1811. Malus discovered that light reflected at a particular angle from plane glass does not possess symmetrical properties of reflection with respect to its direction of transmission. The important discovery of Malus made it possible to explain the earlier observations of asymmetry in connection with double refraction, and to determine that the vibrations in light are transverse to the line of propagation, rather than longitudinal, as in the case of sound.

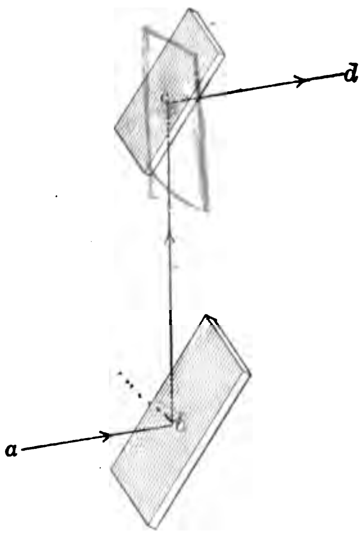


Fig. 217

When a ray of common light *ab* (Fig. 217) falls on a plate of plane glass at an angle of incidence of about 57°, it is reflected in the direction *bc*, in accordance with the laws of ordinary reflection. The ray *bc* is mostly polarized by reflection.

If this polarized ray is incident on a second plate of glass parallel to the first one, the plane of incidence for the upper mirror coincides with that of the lower, and the ray is reflected like ordinary light. If, however, one turns the upper mirror around *bc* as an axis, keeping the angle of incidence on the upper mirror at 57°, the two mirrors will no longer be parallel, and the plane of reflection of the upper mirror will revolve with the mirror. When the upper mirror is thus turned out of parallelism with the lower, the intensity of the light reflected from the second mirror decreases continuously, until the angle of rotation is 90°, when the ray *bc* ceases to be longer reflected. Turning the upper mirror still further, *bc* is again reflected and increases in intensity up to 180°, when the two planes of incidence again coincide. By further rotation of the upper mirror, the reflected light again becomes fainter and fails at 270°. The light reflected from the lower mirror at an angle of about 57° is ~~polarized, for~~ ~~*bc* shows~~ a two-sidedness with respect to reflection.

333. Vibrations in Light Transverse. — If the vibrations in light were longitudinal, as in sound, it is not conceivable that a ray of light could be modified in such a way that it would not be equally susceptible of reflection in all directions. On the other hand, if the vibrations are *transverse*, they may then be reduced to a single plane containing the ray, and will thus be asymmetrical about its direction of transmission. Now reflection from glass is one of the means of reducing ether vibrations to a single plane, and light thus modified is called *plane polarized light.*

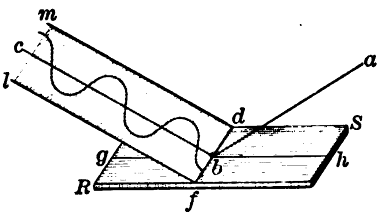


Fig. 218

In light polarized by reflection the vibrations are parallel to the reflecting surface. When, for example, a ray of light *ab* (Fig. 218) is incident on a plate of glass at an

angle of 57°, the reflected ray *bc* is polarized and the ether vibrations are parallel to *fd* and at right angles to the plane of incidence containing *ab* and *bc*.

When such a plane polarized pencil of light *cb* is incident on a glass mirror *RS*, whose plane is parallel to the ether vibrations, a portion of the pencil is transmitted through the glass without change in its plane of vibration, and another portion is always reflected in the direction *ba* with its vibrations unchanged and parallel to *fd*.

334. The Polarizing Angle. — The degree of polarization of a reflected pencil of light varies with the angle of incidence. The angle of incidence for which the polarization is greatest is called the *polarizing angle*. It is different for different substances and indeed for different colors. The following simple law relating to the polarizing angle was discovered by Brewster:

The polarizing angle is the angle of incidence for which the reflected and the refracted rays are at right angles.

Thus, in Figure 219, if *bc*, the reflected ray, makes a right angle with *bd*, the refracted ray, the polarization is most nearly complete, and the angle i is the polarizing angle. The plane of reflection is called the *plane of polarization*. The ether vibrations are at right angles to it.

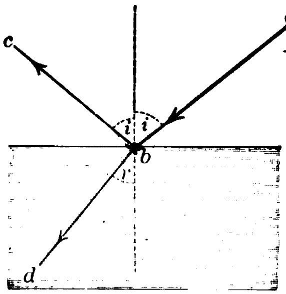


Fig. 219

Since the angle *dbc* equals 90°, so also $i + r = 90°$; consequently $\sin r = \cos i$. Therefore

$$\frac{\sin i}{\sin r} = \frac{\sin i}{\cos i} = \tan i = \mu.$$

Hence the polarizing angle is the angle of incidence whose trigonometrical tangent is equal to the index of refraction of the reflecting substance. Since the different colors have

not the same index of refraction, their polarizing angles for the same substance also differ slightly.

It must not be inferred that for every substance there is an angle of complete polarization. The polarization always increases with the angle of incidence up to a maximum, and then decreases again, after passing the angle of greatest polarization. This maximum is the polarizing angle of the substance. Only a few substances, with a refractive index of about 1.46, polarize light completely by reflection. If the substance is transparent, the refracted ray is also polarized, and in a plane perpendicular to that of the reflected ray.

III. POLARIZATION BY DOUBLE REFRACTION

335. Polarization by Tourmaline. — If a slice of brown or green tourmaline, cut parallel to its optic axis, be held so that a beam of light falls on it normally, the transmitted light will differ in no respect to the eye from the incident beam, except that it is slightly colored by selective absorption in the tourmaline.

But if the transmitted light be examined by means of a piece of plate glass, it will be found to have undergone a remarkable change. While in one direction it is reflected in the same manner as common light, yet when the glass is turned around the beam as an axis, the light varies in intensity; and in one position the reflected light vanishes entirely. The light transmitted by the plate of tourmaline may be reflected in the plane passing through the axis at all angles of incidence; but in a plane at right angles to this, it is imperfectly reflected, and at an angle of incidence of about 57° it is not reflected at all.

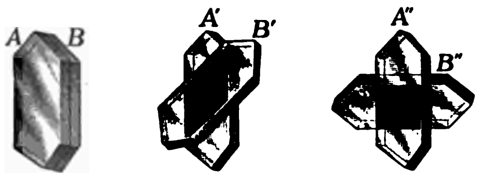


Fig. 220

Further, if a beam transmitted through one plate of tourmaline be examined by a second similar plate, it will be found that in one relative position of the two the light is freely transmitted (Fig. 220, *AB*), while it

becomes feebler and feebler if either plate be turned around in its own plane, as in $A'B'$; and when the two parallel plates have their longer dimensions at right angles, as in $A''B''$, no light whatever passes through them. The light is completely extinguished by crossing two transparent and nearly colorless crystals.

The light transmitted by the tourmaline is *plane polarized.* Tourmaline is a double refracting substance, and it has the property of absorbing the ordinary ray so rapidly that this ray does not emerge at all if the tourmaline plate is from one to two millimeters thick. The first plate of tourmaline is called the *polarizer;* and the plate glass reflector, or the second tourmaline, the *analyzer.*

336. Common Light Compared with Plane Polarized Light. — It is important to observe that the proportion of the beam coming from the polarizer which the analyzer transmits depends on the orientation of the second plate with respect to the first; and that the beam transmitted by the first plate is completely extinguished by the second when the two are crossed. Now, if the vibrations of the ether were longitudinal, it is inconceivable that the crossing of the plates should stop the light, since the rotation of the second plate could not modify longitudinal vibrations. This phenomenon of extinction is therefore held to demonstrate that the vibrations of the ether in light, and in radiation in general, are transverse.

In common light the vibrations are transverse to the direction of propagation and in all planes containing the ray. In other words, the ray is perfectly symmetrical with respect to its direction of transmission. Hence it is that the beam of light incident on the first plate of tourmaline is transmitted equally well in every position of the plate. The rotation of the plate about the ray as an axis produces no change in the intensity of the transmitted light. The action of the tourmaline is to resolve the transverse vibrations of the common

light into simple harmonic components in two directions, one parallel to the optic axis and the other at right angles to it. The former constitute the extraordinary ray, which is transmitted; the latter, the ordinary ray, which is absorbed.

Thus when the ray of common light *IO* is incident on the slice of tourmaline with its faces containing the optic axis *AB* (Fig. 221), the vibrations of the transmitted ray are confined to planes parallel to *AB*, as indicated by the short arrows. When the axis of the second parallel tourmaline plate is inclined to that of the first, its effect is to resolve the incident vibrations parallel to the plane *ABCD* into two components, one in the direction of its axis and constituting the extraordinary ray, which is transmitted, and the other at right angles to its axis and composing the ordinary ray, which is absorbed. When the axes of the two plates are crossed, the vibrations from the polarizer enter the analyzer without resolution, but they form the ordinary ray, which is absorbed. There is then no extraordinary ray, and hence no light gets through the crossed plates.

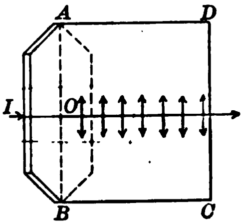


Fig. 221

Again, when the axes of the two plates are parallel, the vibrations from the polarizer are not resolved by the analyzer, but in this case they form the extraordinary ray, which passes through without change.

Common light shows no such evidences of polarization. It is accordingly necessary to assume that, while the motions constituting common light are at right angles to the direction in which the light is traveling, the character of this transverse vibrational form is continually changing, so that in a very short period the vibrations are performed in all azimuths with respect to the ray. On account of this shifting of the plane of vibration, it is impossible to produce interference between two trains of waves from different sources of light, or even from different parts of the same source.

337. Polarization by Iceland Spar. — The two beams of light into which a single one is divided by Iceland spar (§ 328) are both plane polarized. Their polarization may be readily shown by examining them by reflection from plane glass, or by a plate of tourmaline. If the tourmaline is placed in the path of the two rays, with its plane at right angles to them, and if it is slowly turned about the rays as an axis, a position is readily found in which only one image appears on the screen. If then the tourmaline plate is turned further in its own plane, the second image will become visible. It will increase its brightness, while the first one will become fainter and will disappear when the angle of rotation is 90°, the second image only remaining visible.

The positions of the tourmaline when there is only one image show that the ether vibrations of the light forming the ordinary image are perpendicular to a principal plane, while those of the extraordinary image are always in a principal plane. The two rays are therefore polarized in planes at right angles to each other.

338. The Nicol Prism. — A beam of plane polarized light may be obtained by reflection from glass, or by transmission through a plate of tourmaline. By the former method the polarized light is usually mixed with a small amount of common light, and by the latter it is colored by the tourmaline. Any device for getting rid of one of the rays given by Iceland spar leaves a colorless beam of plane polarized light.

The best device for this purpose was invented by Nicol in 1828, and is called the Nicol prism. It is made of Iceland spar in such a way that the ordinary ray is stopped by total internal reflection.

A long rhomb of the spar has its end faces cut so that the natural angle of cleavage is reduced to 68° (Fig. 222, *ACB*). The plane of the figure is a principal plane. The rhomb is cut through along a section perpendicular to a principal plane. The trace of the section is *AB*, which makes

the angle BAC 90°. The two faces of the section are polished and then cemented together by a thin layer of Canada balsam, which has a mean refractive index (1.54) intermediate between those of the ordinary and the extraordinary ray.

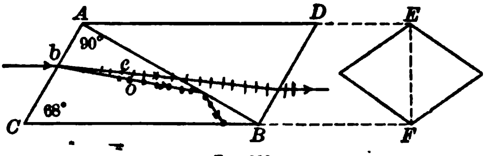


Fig. 222.

When therefore a ray enters the Nicol prism at b, the ordinary ray bo meets the balsam at an angle somewhat greater than the critical angle from the spar to the balsam (68° 15′). It thus suffers total internal reflection and fails to pass through. The extraordinary ray is not totally reflected at the first surface of the balsam because it goes from a medium of lower refractive index for it to one of higher; and it is not reflected at the second surface of the balsam because its angle of incidence is less than the critical angle for the extraordinary ray in the two media. Moreover, since the cemented section is at right angles to a principal plane of the crystal, the vibrations of the ordinary ray are parallel to this section and the ordinary ray is therefore readily reflected by it. Thus the extraordinary ray alone passes through.

The direction of vibration for the ordinary ray is the longer diagonal of the end of the prism; for the transmitted ray it is the shorter diagonal EF. The light emerging from the prism is polarized at right angles to the principal plane.

339. Extinction of Light by Crossed Nicol Prisms.—When the light which has passed through one Nicol prism falls on a second, the amount transmitted depends on the relation of the principal planes of the two. If the shorter diagonals are parallel, the plane polarized light from the polarizer composes the extraordinary ray in the analyzer, and passes on through. But if the analyzer be turned about the beam of light as an axis, the transmitted beam will decrease in brightness, and will disappear entirely when the rotation reaches 90°. The

Nicol prisms are then said to be crossed; the light from the polarizer now forms the ordinary ray for the analyzer, and is lost by internal reflection. In intermediate positions the rectilinear vibrations of the plane polarized light coming through the polarizer are resolved into two rectangular components by the analyzer and in directions corresponding to its two planes of vibration. This resolution takes place in accordance with the usual mechanical law for the resolution of a motion into two components at right angles (§ 32).

340. Effect of interposing Double Refracting Substances. — Let the Nicol prisms be crossed, the direction of vibration for the polarizer being vertical and that for the analyzer horizontal. The field of view is then dark. The introduction between the Nicols of a thin sheet of mica or selenite, with its plane at right angles to the beam of light, in general restores the light on the screen, or to the field of view for direct vision. But if the thin plate is rotated in its own plane, four positions will be found in every revolution in which the interposed plate does not restore the light. These positions are 90° apart. At points 45° from them the transmitted light reaches its greatest intensity.

Mica and selenite are double refracting substances, and the vibrations of the ordinary and of the extraordinary ray are in planes at right angles to each other. Now when the plate is in such a position that its two directions of vibration coincide with those of the polarizer and the analyzer, then the extraordinary ray from the polarizer passes through the crystalline plate without resolution into two components, and it is stopped by the analyzer as the ordinary ray. But for any other position of the mica or selenite plate the case is different.

Assume the mica plate rotated 45° from any position of extinction. Let the full line in Figure 223 *a* be the vibration in the plane polarized light from the polarizer and incident on the mica plate. As soon as this light enters the crystal-

line plate, the single vibration is resolved into two at right angles to each other, *o* and *e*, as represented in *b*. These two trains of waves are transmitted through the mica with

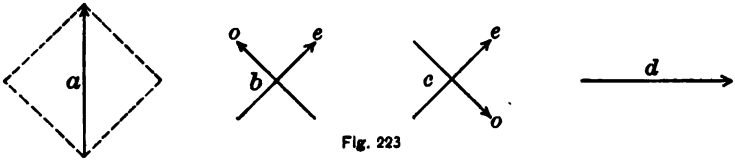


Fig. 223

unequal velocities; and since their frequencies are the same, the one with the slower velocity will gain in phase over the other. Assume the thickness of the plate such that the difference in phase between the two wave trains at the point of emergence is half a period for sodium light, as represented in *c* by the reversal of the arrow *o*. Moreover, since the plate is thin and the beam comparatively broad, the two rays are not visibly separated, and the same portions of the ether are agitated by the two wave trains after emergence from the mica plate. The two vibrations of *c* therefore combine into a single vibration, as represented in *d*.

The light emerging from the mica, compared with the incident beam, is still plane polarized, but its direction of vibration has been rotated through a right angle by the phase difference of half a period introduced by the crystal. Since the analyzer is supposed to be set so as to transmit vibrations in a horizontal plane, it will transmit the light emerging from the mica. In order to extinguish the light now, the analyzer must be turned into a position parallel with the polarizer.

341. Elliptical Polarization.—If the reader will recall §§ 38 and 41 on the composition of simple harmonic motions, it will at once be apparent that the *ellipse* must be the general form of vibration resulting from the resolution of the simple harmonic vibrations of plane polarized light into two components at right angles to each other by transmission through a double refracting body, and the introduction of a phase

difference between the two components by the retardation of the one ray with respect to the other on account of the difference in their velocities of transmission. The ellipse becomes a straight line as one of its limiting forms when the phase difference is a whole number of half wave lengths.

Assuming the relations of the two Nicol prisms and the interposed mica plate to be the same as in the last article, then *a* of Figure 224 is the vibration of the plane polarized

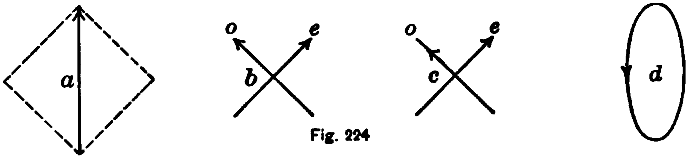


Fig. 224

light incident normally on the mica, *b* the two rectangular components of this vibration immediately after entering the mica, and *c* the two as the light leaves the mica, its thickness being such that the one ray *o* has fallen behind the other one *e* by one eighth of a wave length, as indicated by the head of the arrow in *o*. These two simple harmonic vibrations compound into an ellipse *d*, as may easily be shown by the graphical method of § 41. Hence the light emerging from the mica plate is elliptically polarized. The vibrational form is an ellipse with its plane perpendicular to the direction of transmission.

When this elliptically polarized light enters the analyzing Nicol, the elliptical vibration is resolved into two simple harmonic components, forming the ordinary and extraordinary rays. With a phase difference of one eighth of a period, as represented in Fig. 224, the analyzer will transmit a maximum amount of light in a position parallel with the polarizer, and a minimum when the two Nicols are crossed.

342. Circular Polarization. — The reader who has made himself familiar with the composition of simple harmonic motions at right angles will readily anticipate the possibility of circular polarization of light. Two simple harmonic motions of the same period and at right angles

compound into uniform circular motion when their amplitudes are equal and their phase difference one quarter of a period (§ 41). Now when the mica plate is set so that its directions of vibration are at an angular distance of 45° from that of the incident plane polarized light, then the two components into which the plane polarized light is resolved are of equal amplitude, and they have the same period. If, further, the thickness of the prepared mica plate is just enough to introduce a phase difference

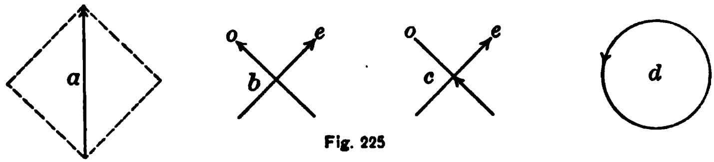


Fig. 225

of one quarter of a period for some selected color, such as yellow sodium light, the emergent beam is circularly polarized (Fig. 225), and the mica is called a quarter wave plate.

In a train of circularly polarized waves the successive ether particles at any instant all lie on the thread of a screw, the pitch of which is the wave length. In right-handed circular polarization (Fig. 226 *R*) the ether particles rotate counter-clockwise; in left-handed polarization they rotate clockwise (Fig. 226 *L*). In *R* the particles *b*, *c*, *d* are moving upward simultaneously and *f*, *g*, *h*, downward; in *L* *b*, *c*, *d* are moving downward and *f*, *g*, *h*, upward, while in both the waves are advancing in the direction of the arrow *A*.

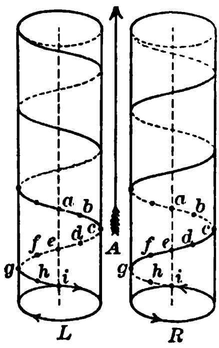


Fig. 226

When circularly polarized light is transmitted through a Nicol prism, there is always an extraordinary ray in the Nicol, with an amplitude of vibration equal to the radius of the circle; therefore no change in the intensity of the transmitted light accompanies the rotation of the prism. If, however, a second quarter wave plate be interposed in the path of circularly polarized light, it will introduce a relative retardation of another quarter wave for the same color, and the emergent light will be plane polarized in an azimuth which may be found by rotating the analyzing Nicol. Two quarter wave plates may thus serve simply to rotate the plane of polarization.

343. Rotation of the Plane of Polarization.—If the two Nicol prisms are set for extinction, and a plate of quartz, cut

perpendicular to its optic axis, is introduced between them, the light will be to some extent restored. It can be extinguished again by rotating the analyzer through a small angle.

Quartz rotates the plane of polarization during the passage of light though it. Some quartz crystals rotate the plane of polarization to the right, looking in the direction of transmission, and these are called right-handed; others rotate the plane of polarization to the left, and are called left-handed.

The rotatory power of quartz is intimately related to its crystalline structure. Fused quartz is not double refracting and does not produce rotation. Further, in some cases when a right-handed and a left-handed quartz crystal are placed side by side, they are found to be symmetrical with respect to a plane between them, like an object and its image in a plane mirror. In other words, they are right-handed or left-handed crystals, corresponding to their power of rotating the plane of polarization to the right or to the left. In many quartz crystals the secondary planes or facets, which distinguish them as right-handed or left-handed, are lacking; but even then their contrasting crystalline structure may be demonstrated by etching.

344. Rotation of the Plane of Polarization Explained.—In quartz the double refraction does not vanish even along the optic axis. The two wave surfaces are therefore not tangent (§ 330), but the spheroid lies wholly within the sphere without touching. Further, both rays in the direction of the axis are circularly polarized, the rotations of the two components being in opposite directions around equal circles. When these circularly polarized rays emerge from the crystal, they recombine into plane polarized light, but with rotation of the plane of polarization.

Let the two uniform circular motions o and e (Fig. 227 A) be resolved into two linear displacements at right angles to each other. They are symmetrical with respect to the line P; the components b and b' are equal and opposite, while a and a' are equal and in the same direction. Hence, if the two uniform circular motions are simultaneously impressed on

the same particle, its velocity parallel to the line P at any instant will be double that of o or e, and perpendicular to it will be zero. The particle will therefore describe simple harmonic motion along the line P with an amplitude twice the radius of the circle.

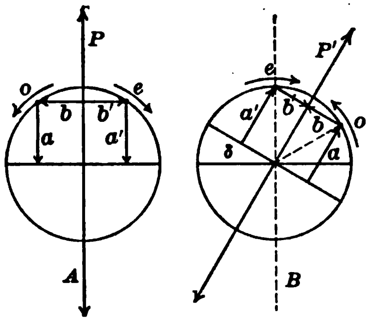


Fig 227

Conversely, the simple harmonic motion P is resolvable into two equal uniform circular motions of the same period as that of P and of equal amplitudes.

Plane polarized light incident normally on a plate of quartz, cut at right angles to its optic axis, is thus resolved into two circularly polarized rays; and if the velocity of transmission of the two through the quartz were the same, the plane of polarization of the light emerging into the air would be parallel to that of the incident ray. But the circular component o, constituting the ordinary ray, travels faster through the quartz than the component e, while the two still have the same frequency. Hence the plane of vibration of the emergent light is rotated in the direction of the circular motion of the component which rotates through the greater angle in traversing the quartz plate. This is the component e, which has the smaller velocity of transmission. Its wave length is shorter than the other, and it rotates through a larger angle in traversing the same thickness of the crystal.

Thus, in Fig. 227 B, o and e are the relative angular positions of the two circular motions as they emerge from the quartz. The component e has gained *in phase* as compared with o, and the two are symmetrical with respect to the line P', the perpendicular displacements, b and b', annulling each other, while the parallel displacements, a and a', are additive. The line P' is then the direction of vibration of the

emergent light. The plane of polarization has been rotated through the angle δ, which is half the angular retardation of the one ray with respect to the other.

The rotation at 20° C. for different wave lengths produced by a plate of quartz one millimeter thick is as follows:

A	*B*	*C*	*D*	*E*	*F*	*G*
12.67°	15.75°	17.32°	21.70°	27.54°	32.77°	42.60°

The angular rotation is nearly inversely as the square of the wave length. This is known as Biot's law.

345. Rotatory Power of Liquids. — It was discovered by Biot in 1815 that some liquids also possess the same power of rotation as quartz, but to a much smaller degree. Of these, solutions of cane sugar have received the most attention, for the reason that the commercial test for the percentage of sugar present is the rotation produced by a column of sugar solution of fixed length and at a definite temperature. This method of quantitative estimation applied to sugar is commonly known as *saccharimetry*.

When heated in solution with dilute acids, cane sugar takes up water and splits into two molecules. The two have the same chemical composition, but one rotates the plane of polarization to the right, and is therefore called *dextrose;* the other rotates it to the left, and is called *levulose*. The levulose rotates the plane of polarization more strongly than the dextrose, and the mixture therefore rotates to the left, or the rotatory action of the cane sugar has been inverted. Whence the name "invert sugar."

IV. COLORS BY POLARIZED LIGHT

346. Colors produced by an Interposed Film. — When the polarizer and the analyzer are set so that the latter quenches the light transmitted by the former (§ 340), a lamina of mica or selenite, held obliquely in the path of the polarized white light between the two Nicol prisms, not only brings

light into the dark field, but shows brilliant colors. The restoration of the light has already been explained; it remains now to account for the colors.

Let us suppose that the selenite film (which gives more brilliant colors than mica) is just thick enough to introduce a difference of phase between the two component vibrations in the selenite of half a period for the longest waves of red. Then, since the short waves of extreme violet are about one half the length of those of extreme red, the selenite lamina will produce a phase difference of a whole period for the violet. The two components for the latter will thus emerge from the selenite without any phase difference, and will recombine into plane vibrations in the same direction as those of the polarized light incident on the lamina. The red, however, will emerge from the selenite with one of its component vibrations half a period out of phase with the other, and the two will combine into plane vibrations at right angles to those of the incident beam, as represented in Fig. 223 *d*. The plane of vibration for red has been rotated in this manner through a quarter turn, and it will form the extraordinary ray for the analyzer and will pass through. The violet, on the other hand, is polarized in a plane at right angles to that of the red and will suffer extinction by the analyzer. If then only red and violet light were mingled in the plane polarized beam incident on the selenite, the analyzing Nicol would cut out all the violet and transmit all the red; but if the analyzer were rotated through a quarter turn, it would cut out all the red and transmit all the violet, and the continued rotation of the analyzer would allow the red and the violet to be transmitted alternately.

If, now, the incident light is white, all wave lengths between those for red and for violet are present, and the vibrations for these after passing the selenite are ellipses of various forms. The vibrational forms for the orange and yellow near the red are ellipses approaching straight lines, and their components parallel with those for red are much larger than

those parallel with the vibrations for violet. The orange and the yellow will therefore be largely transmitted along with the red; while those nearest violet, namely, the blues and shorter greens, will be largely transmitted with the violet. It follows that the light transmitted by the analyzer when crossed with the polarizer will be some shade of red; when parallel with the polarizer, some shade of blue. The colors transmitted in the two positions make up the whole of the white light incident on the selenite, and they are therefore complementary to each other.

A double refracting lamina thick enough to produce a phase difference of many periods for any color will also produce a phase difference of many periods for a number of other colors distributed throughout the spectrum. These colors may then be cut out by the analyzer, and the remaining transmitted colors will also be so widely distributed throughout the entire range of the spectrum that they will together reproduce the effect of white light. Hence, colors are produced in polarized light only when the double refracting lamina gives rise to a phase difference of only a few periods.

347. Colors Due to a Plate cut at Right Angles to the Optic Axis. — Extremely beautiful effects are produced by plates of uniaxial crystals *cut perpendicular to the optic axis*, such as a section of Iceland spar, in a beam of *converging* plane polarized light. The plate is placed between the two Nicol prisms in a converging beam so that the central ray of the converging cone of light is normal to the plate. This central portion of the cone of light passes along the optic axis of the section without undergoing double refraction, but all the other rays traverse the crystal section obliquely and are doubly refracted. The further any ray is from the axis of the cone, the greater is the obliquity of its path through the plate and the greater the thickness traversed. Also, the more oblique the ray, the greater the difference in the velocity of transmission of the two component disturbances. Since at the same distance from the optic axis each of these two causes of phase difference has a constant value, it follows that there must be the

same phase difference for any one color at all points of a circle conceived as drawn on the screen around the axial ray of the cone of light as a center. Hence, a system of concentric rings appears on the screen in iridescent colors like those of Newton's rings (§ 302).

When the polarizer and analyzer are crossed, the colored rings are traversed by a black cross (Fig. 228). This cross is explained as follows: Since the optic axis of the plate of spar is perpendicular to its surface, every diameter of the system of rings is the trace of a principal plane. The vibrations of the ordinary ray are normal to a principal plane and therefore tangential to all the concentric circles; those of the extraordinary ray are in a principal plane, or radially in the circles. Hence, along the two diameters representing the planes of vibration of the polarizer and the analyzer, the directions of vibration for the interposed plate are the same as those of the polarizer and the analyzer. In these two directions, therefore, the thin plate does not resolve a ray into two components, and since the Nicols are crossed, the field is dark. Along all other diameters, the tangential and radial directions of vibration for the plate are inclined to those of the polarizer and the analyzer; therefore double refraction takes place with colors as already explained.

Fig. 228

If the analyzer is turned so as to be parallel with the polarizer, a white cross takes the place of the black one. The colored rings are then projected on a bright field as a background.

Similar dark and bright crosses are obtained from artificial crystals of salicin on glass. The crystallization starts from many centers, and around them grow crystals with a radial structure and circular in outline. The maximum resolution of the plane polarized light occurs in each

circle midway between the planes of vibration of the polarizer and the analyzer, while in these latter directions there is no resolution of the polarized light. Along these lines therefore the field remains dark, giving the black crosses (Fig. 229). When the analyzer is turned through 90°, the black crosses change into white ones, and all the little crystals appear to revolve with the analyzer.

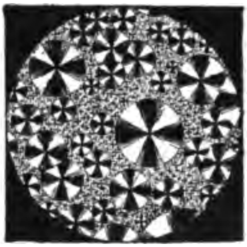
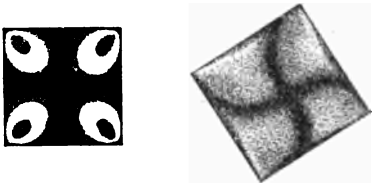


Fig. 229

348. Colors Produced by Stresses in Glass. — Well-annealed glass does not show double refraction; but it may be made to acquire this property by unequal stresses in different directions. It is then no longer an isotropic body. Such stresses are evoked in glass by sudden chilling after heating. When glass so treated is traversed by plane polarized light, chromatic phenomena are obtained analogous to those due to double refracting crystals. A considerable variety may be given to the colored curves by varying the form of the glass plates, making them circular, square, rectangular, or triangular in pattern.



A *B*

Fig. 230

The introduction of one of these plates between the crossed Nicol prisms allows light to pass through, and colors spring forth of surprising brilliancy. Figure 230 *A* represents the black cross appearing on the dark field with a square of unannealed glass, and Figure 230 *B*, the figure when the plate is turned around in its own plane.

Fig. 231

Similar effects are produced by the mechanical compression of annealed glass. When a square piece of thick plate glass is compressed by turning a screw with the thumb and finger (Fig. 231), the light appears in the field where the pressure is applied at opposite points of the block, and it extends farther into the block

as the pressure is increased, showing distortion through the whole body.

The double refraction produced by a bending stress applied to a long polished strip of thick glass by means of a small hand press (Fig. 232) is very instructive. The polarized light passes through at right angles to the bending stress. The glass restores the light on both edges, but leaves a dark band through the middle from end to end. The concave surface of the bent strip is compressed, and the convex surface is stretched. The middle is the neutral axis, which is neither compressed nor elongated.

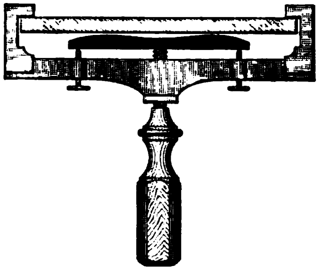


Fig. 232

Another very instructive experiment in double refraction in glass under stress was devised by Biot as long ago as 1820. It demonstrates the alternate compression and extension of a strip of glass at a node when vibrating longitudinally (§§ 199 and 249). A glass tube serves as well as a strip of glass and is more convenient to handle. It should be placed obliquely across the field of polarized light between the two Nicols and should be held firmly by its middle point. This point is a node, and the light should pass through the glass as near the node as practicable. If the tube is inclined to the planes of vibration of the crossed Nicol prisms, the field remains dark until the tube is thrown into vigorous vibration by rubbing it lengthwise with a moist cloth. Then every time a note is produced a band of light flashes across the field of view. The distortion of the tube due to the passage along it of compressional waves renders it double refracting.

Since the tube is alternately compressed and stretched, it is obvious that twice during every complete vibration it is in its original isotropic state; therefore the illumination of the field must be interrupted by short periods of darkness. The change occurs too rapidly to be observed directly; but Kundt verified this conclusion by observing the light in a rotating mirror. The elongated band of light seen in the mirror was crossed by dark spaces. Moreover, Kundt determined that the glass is optically positive like quartz while it is extended, and optically negative like Iceland spar while it is compressed.

CHAPTER XII

OPTICAL INSTRUMENTS

I. THE EYE

349. The Photographic Camera. — The essential parts of a photographic camera are a dark chamber, or "camera obscura," and an achromatic converging lens provided with a screen on which to form a real image. In Figure 233, *BC* is the dark chamber, which is blackened inside and usually adjustable in length, *LL'* is the converging lens combination, and *E* the screen or sensitized plate.

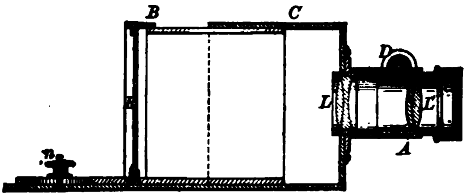


Fig. 233

In front of the lens is a shutter and, if the lens is single, a diaphragm with small openings. This cuts off the marginal rays and reduces spherical aberration. It also gives greater depth of focus for landscape photography, so that objects at different distances may be approximately in focus at the same time, analogous to the pinhole camera (§ 257), in which the focus is nearly independent of the distance of the object.

The requirements of a good photographic lens are quite different from those of instruments used for visual impressions. As the photographic films in common use are most sensitive for the shortest waves, the lens must be achromatically adjusted around the blue-violet as a mean instead of the yellow-green, which is brightest to the eye. The photographic camera is expected to form an image covering a wide area which must be flat and free from distortion, and all points of it must

be focused together. Distortions cannot be entirely eliminated by a diaphragm and a single lens. With two lenses or sets of lenses and a diaphragm between them, the distortions correct each other. Such a combination is called a *rectilinear doublet.*

350. The Eye. — The eye resembles a small camera, into which light enters only through a lens. In the camera the lens forms an inverted image on the sensitized plate, where the light initiates a chemical change in the sensitizing silver salts; the lens of the eye forms an inverted image on the retina and agitates its sensitive nerve terminals.

Figure 234 is a vertical section through the axis of the eye. The *sclerotic coat H* is a tough, fibrous substance, known as "the white of the eye." It is opaque except in front, where it becomes a transparent coat *A*, called the *cornea.* Behind the cornea is the *iris D*, a colored curtain perforated by a circular aperture *C* called the *pupil.* The pupil adjusts itself in size automatically with reference to the intensity of the light. Between the iris and the *crystalline lens E* is a transparent fluid called the *aqueous humor.* The chamber behind the lens is filled with a thin albuminous fluid called the *vitreous humor.* The *choroid coat*, filled with a black pigment to prevent internal reflection, lines the walls of this chamber; on it is spread the *retina*, which is traversed by a network of nerves branching from the *optic nerve M.*

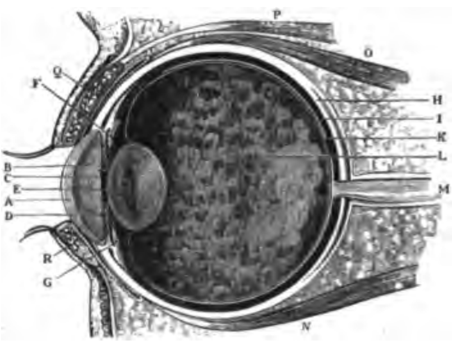


Fig. 234

The lens is built up of transparent horny layers, which increase in refractive index toward the center. This increase in the index of refraction toward the axis serves for the partial correction of spherical aberration (§ 293), which is

further diminished by the iris diaphragm, as in the photographic camera.

In the camera the distance between the lens and the plate is adjustable for objects at different distances. In the eye the corresponding distance is fixed. When the eye is at rest, or meditatively fixed on space, it is adjusted for vision for infinitely distant objects. To look at nearer objects requires an effort of *accommodation*. Muscles attached to the periphery of the lens change its curvature by contraction or relaxation, and thus enable it to focus on the retina either very distant or very near objects. The range of accommodation is the same as if the eye were provided with a series of lenses of all focal lengths for objects ranging in distance from infinity to about twenty centimeters. This power of accommodation wanes with advancing age, so that the image for near objects falls behind the retina. Hence the necessity for converging reading glasses.

351. The Blind Spot. — The point in the retina where the optic nerve enters it is not sensitive to light; it is accordingly called the *blind spot*. When the image falls on this spot, there is no visual impression produced. This can easily be demonstrated by the aid of Figure 235. Hold the

Fig. 235

book with the circle opposite the right eye. Then close the left eye and turn the right to look at the cross. Move the book toward the eye from a distance of a little more than 30 cm. (1 foot), and a position may readily be found where the black circle will disappear. Its image then falls on the blind spot. It may be brought into view again by moving the book either nearer the eye or farther away.

352. Defects of the Eye. — A normal eye in its passive condition focuses parallel rays on the retina; but many eyes are not normal and have defects of several sorts. Those of most

frequent occurrence are near-sightedness, far-sightedness, and astigmatism.

If the eye when relaxed focuses parallel rays in front of the retina, it is *near-sighted.* The length of the eyeball from front to back is then too great for the focal length of the crystalline lens. The correction consists in placing in front of the eye a concave lens that makes with the lens of the eye a less convergent system than the crystalline lens itself.

Let d be the greatest distance for distinct vision of a near-sighted person. Then if the focal length of the concave lens is d, and if it is held close to the eye, parallel rays from a distant object will enter the eye as if they came from the principal focal point of the lens, which is at the greatest distance for distinct vision d (Fig. 236). Provided with such

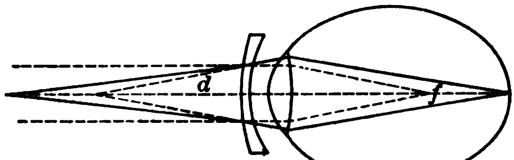


Fig. 236

near sighted

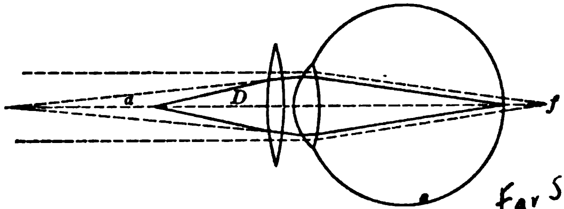
a lens and possessed of the power of accommodation, the near-sighted eye may then have distinct vision for objects at all distances.

If the eye when relaxed focuses parallel rays from a distant object behind the retina, it is *far-sighted.* The length of the eyeball is then too small to correspond with the focal length of the crystalline lens. The correction for far-sightedness consists in placing in front of the eye a converging lens, making with the lens of the eye a more converging system than the crystalline lens alone.

In this case let d be the least distance for distinct vision, and D that of normal distinct vision for small objects like the letters on this page. It is customary to take D as 25 cm.

(10 in.). Then the focal length of the correcting lens may be found by substitution in formula (64). Since d and D are both positive, $1/D > 1/f$, or D is less than f. The converging lens then gives a virtual image, and light from an object at a distance D enters the eye as if it came from a distance d (Fig. 237).

Sometimes the front of the cornea has different curvatures in different planes through the axis; that is, it has a somewhat cylindrical form. Horizontal and vertical objects at



Far Sighted

Fig. 237

the same distance are not then in focus at the same time. This defect is known as *astigmatism*. It is corrected by the use of a lens one surface of which at least is not spherical, but differs from it in the opposite sense to that of the defective eye. The astigmatism of the two eyes is not usually the same.

353. Binocular Vision. — For distances in excess of a few hundred feet there is little difference between vision with one eye and with both if the illumination is good. But for near objects binocular vision introduces a new element. It is easy to demonstrate that when both eyes are directed to the nearer side of an object, the images of a point farther away do not fall on corresponding retinal points, and hence there is double vision. Hold two lead pencils vertically in front of the eyes with an interval of about six inches between them. When the eyes are fixed on the nearer pencil, the farther one is seen double. Closing the right eye causes the right image to disappear, and closing the left eye, the left

image. If the gaze is fixed on the more distant pencil, the nearer one is seen double, but now the right image is given by the left eye and the left image by the right eye.

The perception of depth in binocular vision is attained by perfect fusion of the images of a point at one distance from the eyes, while there is imperfect fusion of the images of points more remote and less remote. One eye occupies a position sensibly different from that of the other, and the retinal images are therefore not identical. On the whole, the impression received is that of a single object; but the right eye sees more of the right side of this object, and the left eye more of the left. Binocular vision thus comprehends more than vision with one eye. The two images are unconsciously combined into a single impression embracing the features of both retinal pictures. The effect of this fusion is the perception of depth, and the observer is as unconscious of double images as he is of the inversion of all retinal images.

Pictures to be viewed with binocular vision in a stereoscope are taken with a camera provided with two lenses, giving dissimilar images like those of the two eyes. A print from one of these negatives is viewed with the right eye and one from the other with the left eye only. The fusion of the two impressions gives the effect of relief in the picture.

354. Irradiation. — The increase in the apparent size of an object as it becomes more highly luminous is known as *irradiation*. Thus, the filament of an incandescent lamp appears to become thicker as it passes from a red-hot to a white-hot temperature; the crescent of the new moon appears to belong to a larger circle than the remainder of the disk, which is only faintly illuminated by light reflected from the earth; a candle or gas flame appears to be continuous, though the incandescent particles of carbon are not in contact with one another. Strongly illuminated white objects, or those of a very bright color, appear larger against a dark background

than they really are. Irradiation probably arises from the fact that the impression produced on the retina extends beyond the outlines of the geometrical image.

355. Persistence of Visual Impressions. — When a piece of ignited charcoal is rapidly whirled about in a circle, the appearance produced is a circle of fire. The spokes of a rapidly rotating wheel cannot be seen as separate images, but they blend into one another and give the impression of a translucent disk. Drops of falling rain have the appearance of liquid threads. The roughened part of a small vertical stream of water looks continuous below the smooth portion, while in reality it consists of distinct drops. An electric arc lamp, fed with alternating current, appears to give a continuous light if the frequency is above about twenty-five cycles a second; but when the arc is photographed on a rapidly falling plate, it is found to be extinguished with every reversal of the current, and is really discontinuous. If it is viewed at night with a sudden turn of the head and eyes, the discontinuity may be seen by catching several stages of the illumination on different portions of the retina.

The apparent continuity in all such cases is due to the fact that visual impressions persist after the external stimulus ceases. The explanation is physiological rather than physical, but the persistence of impressions made on the eye is involved in so many physical phenomena, such as manometric flames (§ 238) and Lissajous's figures (§ 248), that it cannot be ignored in physics. The duration of a visual impression depends on the sensitiveness of the retina and the intensity of the light. Plateau found an average duration of half a second.

II. MICROSCOPES AND TELESCOPES

356. The Simple Microscope. — A converging lens of rather short focal length may be used as a *simple microscope* or magnifying glass. The object is usually placed just within the principal focal distance; the image is then virtual, erect, and enlarged. The image *ab* (Fig. 238) subtends at the center of the lens the same angle as the object *AB*; and since the

eye is placed close to the lens, the angles subtended by the image at the center of the lens and at the eye are nearly the same. The image is usually brought to the distance of normal distinct vision for the eye, say 25 cm. The magnification is then approximately $25/f$, where f is the focal length of the lens in cm.

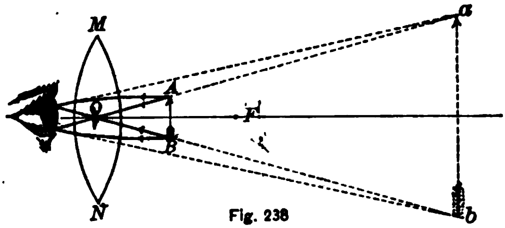


Fig. 238

The value of a simple magnifier consists in the ability it offers of bringing the object much nearer the eye than is possible without it, and thus in effect increasing the size of the retinal image. The lens may be considered as merely extending the power of accommodation of the eye.

357. The Compound Microscope. — The linear magnification of a simple magnifier cannot be extended much beyond one hundred diameters. To obtain still higher magnifying powers a combination of converging lenses, known as a compound microscope, is used. In its simplest form it consists of an object lens or train of lenses O (Fig. 239), which forms a magnified real image ab, and an eyepiece O' used as a simple magnifier with ab as the object. The eyepiece gives a virtual image $A'B'$. Magnification is produced both by the objective and the eyepiece. The former is of short focal length and the object is placed just beyond its principal focus, so that a real and inverted image is formed very near the focus of the eyepiece. The eyepiece as a simple microscope may give a virtual image at a distance from the eye of normal distinct vision; but since the normal eye at rest is adjusted

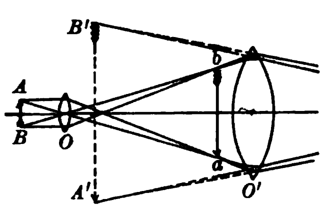


Fig. 239

for objects at a great distance, it is easier for the eye if an optical instrument is so arranged that the rays entering the eye are nearly parallel. This can easily be done with the compound microscope by placing the eyepiece so that the image *ab* coincides with its principal focus. The focusing is then done by the instrument instead of using the accommodation of the eye.

The above is only an outline of the principle of the microscope. The reader is referred to special treatises for the complete theory of the modern microscope. In it both the objective and the eyepiece consist of a combination of lenses, the former sometimes containing as many as ten separate ones to correct for spherical and chromatic aberration and to give a flat focal surface.

358. The Astronomical Telescope. — In the astronomical telescope, which gives an inverted image, the objective forms a real image which is viewed with the eyepiece or ocular, as in the compound microscope (Fig. 240). The employment of

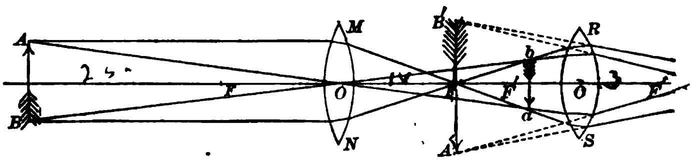


Fig. 240

a large objective is for the purpose of collecting enough light to permit of large magnification without too much loss in brightness.

Since the opposite angles between the two secondary axes through *O* are equal, the angular magnitudes of the object and its real image *ab*, seen from *O*, are the same; but if the eye approaches *ab* to the least distance of distinct vision, the image will increase in magnitude in the ratio of the distances of the objective and the eye from the image. If *F* is the focal length of the objective in centimeters, the magnification would be $\frac{F}{25}$. But the magnification is increased by the

eyepiece by the factor $\frac{25}{f}$, where f is the focal length of the eyepiece (§ 356). The total magnification then becomes $\frac{F}{25} \times \frac{25}{f} = \frac{F}{f}$, or the focal length of the objective divided by that of the eyepiece.

If the telescope is turned toward a bright sky, the eyepiece may be focused on the illuminated objective, and an image of it, called the *ocular circle*, may be received on a piece of white paper very near the eyepiece. Now, the size of the objective and that of its image are directly proportional to their respective distances from the eyepiece, $F+f$ and a (a the distance of the real image from the eyepiece); also equation (64) may be written $\frac{1}{F+f} + \frac{1}{a} = \frac{1}{f}$. Multiplying both sides of this equation by $F+f$ and subtracting unity, we have $\frac{F+f}{a} = \frac{F}{f}$. But $\frac{F}{f}$ is the expression for the magnification. It follows that the magnification is equal to the ratio of the diameter of the objective to that of the ocular circle.

The ocular circle is the smallest area traversed by the light after it leaves the telescope, and it marks the best position for the eye of the observer. To facilitate the placing of the eye in this position, a brass diaphragm, with a hole in its center, is screwed into the eye end of the telescope; the proper place for the eye is close to this hole. If the pupil is smaller than the ocular circle, there is loss of light. Now the smaller the ocular circle, the larger the magnifying power of the telescope; hence, the *lowest* power that can be used to advantage is one that makes the ocular circle no larger than the pupil of the eye. This latter in feeble light may be assumed to be from two fifths to one half a centimeter in diameter. It follows that the lowest power which can be used to advantage is from two to two and a half times the diameter of the objective in centimeters.

359. Galileo's Telescope. — *Galileo's telescope* is the simplest of all telescopes giving an erect image. To the astronomical telescope an inverting system must be added if the image is to be erect. If an ocular system, similar to a compound

microscope, be substituted for the ordinary eyepiece of an astronomical telescope, the image will be erect. But this substitution means more surfaces and more loss of light.

Galileo's telescope consists of only two lenses, the objective *MN*, and the diverging system *RS* (Fig. 241). The distance between the two lenses is approximately the difference of their focal lengths; in the astronomical telescope it is their sum. It was invented by Galileo in 1609, and about six months later he discovered with it four satellites of Jupiter, and soon thereafter the mountains on the moon, spots on the sun, and the variable phases of the planet Venus.

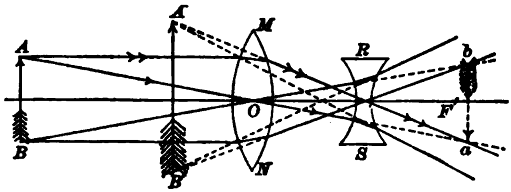


Fig. 241

In consequence of the divergence of the rays emerging from the eyepiece, Galileo's telescope has only a small field of view, and the greater magnification, the smaller the field. For this reason the use of this type of telescope is restricted to cases where only moderate magnification is required. The binocular *opera glass* consists of a pair of Galilean telescopes combined with their axes parallel. It produces an image in each eye and thus secures greater brightness. The magnification in opera glasses is rarely more than three diameters.

A peculiarity of this telescope, which precludes its use for astronomical measurements, is that no cross wires can be used in it. To be of service a cross wire must coincide with the real image given by the objective; but no such image is formed in Galileo's telescope. Another peculiarity is the absence of the ocular circle. There can be none because the image of the object glass formed by the eyepiece is virtual.

Problems

1. The frequency n for the spectral line H_1 is 756×10^{12}. What velocity of a star in the line of sight is derived from a displacement of H_1 equal to 0.15 of a unit toward the violet end of the spectrum?

2. Since light travels in straight lines, the intensity of illumination is inversely as the square of the distance. If an incandescent light and an arc light of 16 and 400 candle power respectively are placed 180 cm. apart, where must a white screen be placed to receive equal illumination from the two sources?

3. If a crystal of Iceland spar has its faces cut parallel to its optic axis and 4 cm. apart, how far below the upper surface are the ordinary and extraordinary images of a mark on its lower face when viewed normally?

4. A narrow aperture 0.05 mm. wide is placed parallel to a screen 3 m. distant. If the aperture be illuminated with a beam of parallel rays of sodium light ($\lambda = 5890 \times 10^{-10}$ m.), what will be the distance between the central image and the first dark band of the diffraction image on the screen?

5. A glass grating is ruled with 4245 lines to the centimeter. When plane waves of sodium light are incident on this grating normally, the image of the second order is at an angular deviation of 30° (θ_2). Find the wave length.

6. If the greatest distance for distinct vision for a near-sighted person is 10 cm., what should be the focal length of glasses to read at a distance of 25 cm.?

7. If the nearest distance for distinct vision for a far-sighted person is 80 cm., what should be the focal length of converging lenses to read at a distance of 25 cm.?

HEAT

CHAPTER XIII

NATURE AND EFFECTS OF HEAT

I. HEAT AND TEMPERATURE

360. Sensations of Heat and Cold. — If one grasps an iron bar, heated in a blacksmith's forge, it is said to feel *hot;* on the other hand, if one holds in one's hands a block of ice, it is said to feel *cold*. The words *hot* and *cold* belong to primitive language and express sensations with which every one is familiar. They are not given by sight or hearing, but by a specialized sense which is often confused with that of touch. The iron bar feels hot because it imparts *heat* to the hand; the ice feels cold because the hand loses *heat* to it.

Experience has thus made us acquainted with a physical state of bodies which is quite independent of their mass motions or the forces acting on them. This physical condition is made known to us in common experience by the sense of heat, and the experience is often coupled with painful consequences. The sensation and its consequences are associated with fire and they are ascribed to *heat* as the cause. It is with heat as a physical quantity that we have now to deal.

361. Heat a Form of Energy. — Up to the beginning of the last century heat was thought to be a very tenuous form of matter, or subtle fluid, that could penetrate liquid and solid bodies with ease. This imaginary fluid was called *caloric;* hence the caloric theory of heat.

About the beginning of the nineteenth century, the classical experiments of Count Rumford and Sir Humphry Davy demonstrated that the caloric theory was no longer tenable. The former observed that in the boring of brass cannon with a blunt drill in the arsenal at Munich much heat was generated, while only a small quantity of metal was abraded. This abraded material showed no change in its capacity for heat, and at the same time no stage was reached where the heat of the metal showed any signs of exhaustion. All that was required for the continued generation of heat was the further expenditure of work in turning the drill against friction. Rumford concluded that the heat which a system of bodies can thus continue to furnish indefinitely cannot be a material substance.

About the same time Davy, then a young man of only twenty-one, showed by experiment that enough heat could be developed by the friction of one block of ice against another to melt the ice. At the same time the water, instead of having its heat capacity diminished by the process, as was claimed by the calorists in the case of the brass abraded in boring the cannon, has actually more than twice the capacity for heat that the ice has from which it comes.

Near the middle of the last century Joule of Manchester in England demonstrated by extensive experiments that a definite amount of mechanical work is the equivalent of a definite amount of heat. It was thus finally settled that there is no such substance as caloric, but that *heat is a form of energy*.

If the work done on a system of bodies is not directed by the environment or the mechanism into some other form of energy, it always appears as an equivalent of heat. Thus, an expert blacksmith may heat a small bar of iron red-hot by rapid dextrous hammering; the kinetic energy of a swiftly moving cannon ball is converted into heat when it strikes the target, as evinced by the flash of light; the friction of a piece of steel on an emery wheel converts the

energy expended in turning it into sufficient heat to raise the temperature of the abraded particles to the point of ignition; the kinetic energy of a heavy moving railway train is converted into heat chiefly by the friction of the brakes. Glowing bits of metal from the wheels often give evidence of the high temperature of the abraded iron. The heating of a compression pump, when used for the inflation of a bicycle or an automobile tire, bears testimony to the conversion of mechanical work into heat by the compression of air. The converse expenditure of heat in doing mechanical work by means of a heat engine is now so common an operation that comment relating to its bearing on the question of the nature of heat as a form of energy is superfluous. In the case of high-speed machinery, it is difficult to prevent the reversion of this energy into wasteful heat at the bearings.

We conclude therefore that heat is a form of energy; and, if so, it can be measured in ergs, joules, kilogram meters, or foot pounds.

362. The Molecular Theory of Heat. — The kinetic theory of matter, from which most of the properties of a gas have been deduced, assumes that the molecules of a body are in a state of incessant agitation (§ 164). The innumerable collisions between molecules, and their mutual jostling, after disturbance from without, must speedily produce a uniform state, in which the molecular motions are in all possible directions.

After this steady state has been established, the molecular energy is uniformly distributed throughout the body; it is largely kinetic, but may be in part potential. Heat then is supposed to be the energy due to the irregular motion of the molecules of a body. The caloric theory demanded the conservation of heat as a necessary corollary. It was supposed to be invariable in amount, but might become hidden or "latent." In the modern theory heat is convertible into other forms of energy and is incessantly varying in amount.

The caloric theory is inconsistent with the doctrine of the conservation of energy. Some of its terminology still remains as relics of an obsolete theory of heat.

363. Temperature. — Common observation teaches us that when a hot body is placed in contact with a cold one, the latter becomes warmer and the former cooler. This process, tending toward thermal equilibrium, is assumed to be the passage of heat from the hot body to the cold one, and the assumption is entirely independent of any theory of heat. The two bodies between which there is a passage of heat are said to be at different temperatures. If A is the hot body and B the cold one, the temperature of A is higher than that of B, and the heat flows from the body of higher temperature to the one of lower.

Temperature may be defined as *the thermal state of a body which determines the transfer of heat between it and other bodies.* In the unaided or spontaneous transfer of heat, the body losing heat during the process of equalization is said to be at the higher temperature.

Temperature should not be confused with quantity of heat. A cupful of boiling water is at a higher temperature than a pailful of tepid water, but the latter gives out the greater quantity of heat in cooling down to the freezing point because of the greater quantity of water.

364. Sensation not a Reliable Measure of Relative Temperature. — Incorrect estimates of relative temperatures are often drawn from the sensation of touch. Our judgment of the temperature of the atmosphere is influenced by the wind and by moisture in the air. If the left hand is held in hot water and the right in ice water for a few seconds, and immediately thereafter, if both are thrust into tepid water, the latter will feel cold to the left hand and hot to the right. The sensations are dependent in no small degree on immediately preceding experience. Even scientific instruments for measuring temperature are not entirely independent of their pre-

vious history, but they are much more so than the sensations of the hand.

It is evident that some method independent of physical sensation must be employed for the reliable measurement of the relative temperatures of bodies. The method most commonly used is the increase in the volume of a body attending a rise in its temperature.

II. THERMOMETRY

365. The Thermometer. — An instrument designed to measure temperature, as a physical quantity, is called a *thermometer;* if the temperature to be measured is high, it is called a *pyrometer.*

The choice of the thermometric substance depends largely on the purposes for which the thermometer is to be used. For the construction of a standard scale of comparison, hydrogen gas is the substance most commonly employed. It may be inclosed in a constant volume thermometer, in which the increase of pressure is measured for a given increment of temperature when the gas is inclosed in a vessel of constant volume; or in a constant pressure instrument, in which the increase of volume under constant pressure for a given increment of temperature is measured. In the former it is assumed that the change of pressure is proportional to the change of temperature producing it. In the latter form the same assumption is made relative to the volume. Neither assumption can be demonstrated experimentally with precision, for to do so implies the previous possession of an accurate instrument possessing this same property. But there are theoretical reasons for supposing the law to be correct, at least for gases of sufficient tenuity, which follow Boyle's law most closely.

For domestic and commercial purposes, the thermometric substance in common use is mercury in glass. For temperatures below the freezing point of mercury (§ 370), alcohol is often employed, but on account of its irregularities,

toluene is to be preferred. Pentane remains liquid at a much lower temperature than alcohol, and can be used for extremely low temperatures. All such thermometers must be carefully compared with a standard hydrogen scale.

366. The Mercurial Thermometer.—The *mercurial thermometer* consists of a capillary glass tube of uniform bore, at one end of which is blown a bulb, either spherical or cylindrical (Fig. 242). The bulb and a portion of the tube are filled with mercury. Small changes in the volume of the mercury in the bulb due to changes of temperature show themselves by an appreciable motion of the end of the thread of mercury in the capillary tube.

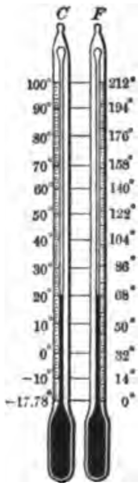


Fig. 242

The observed dilatation of the mercury is its excess above that of the glass envelope. The assumption is that a change of temperature is proportional to the apparent change in volume of the mercury. But thermometers made of different kinds of glass do not quite agree among themselves, and none of them agree precisely with the normal hydrogen scale.

Mercury is chosen as the common thermometric substance for several reasons. Among them is the fact that it can be obtained quite easily in a pure state by distillation in a vacuum; also, its dilatation is relatively large in comparison with that of glass, so that it indicates small changes in temperature; in a pure state it does not adhere to the walls of the glass tube, and the end of the thread of mercury is therefore well defined; its capacity for heat is small, and there is thus little transfer of heat between it and the body whose temperature is to be measured in coming to thermal equilibrium.

On the other hand, the glass is always more or less affected by its previous history. The bulb contracts slowly for a long period after it has been blown in acquiring molecular equilibrium, and this slow contraction raises the readings for all temperatures of the scale. Moreover, the apparent dilatation of the mercury is affected by the irregular dilatation of the glass; besides, on account of the great density of mercury, its hydro-

static pressure enlarges the bulb, and the readings of a sensitive vertical instrument, in which the bulb is large and the stem fine, are not in exact agreement with those in a horizontal position. A sensitive thermometer should, therefore, always be compared with a standard thermometer in the position, horizontal or vertical, in which the former is to be used.

367. Two Fixed Points on the Scale. — If the expansion of mercury by heat is to be utilized in measuring temperature, not only must a scale be chosen, but there must be certain fixed points of reference before the scale can be applied. No two bulbs and their tubes agree in size, and each scale must be graduated separately. This may be done by comparison with a standard, but the graduation is commonly applied first and the correction afterwards.

The fixed points of temperature, which may be verified at any subsequent time, are the melting point of pure ice and the temperature of steam from water boiling under standard atmospheric pressure. They are called the *freezing point* and the *boiling point*. The first is determined by marking on the stem the position of the end of the thread of mercury when the thermometer is surrounded with pounded ice in a room above freezing temperature. It is of prime importance that the ice should be free from contamination such as salt; otherwise the freezing point will be too low.

The determination of the boiling point is much more difficult, for the temperature of steam in contact with boiling water varies with the pressure under which it is formed. The thermometer is placed in the steam and not in the water because under the same pressure the steam is always at the same temperature, while that of the boiling water depends slightly on the vessel and on the presence of foreign substances in it. Also, if the steam escapes freely into the open air, the pressure of the steam is the same as that of the atmosphere, and this may be measured by the barometer. If the reduced height of the barometer is not 760 mm., a correction must be applied to the observed boiling point (§ 405). That two fixed points on the scale of a mercurial thermom-

eter are necessary in order to make the readings of one thermometer comparable with those of another appears to have been first recognized by Newton in 1701.

368. Thermometric Scales. — Fahrenheit appears to have made his first thermometers about 1714, but the earliest published description of them he contributed to the *Philosophical Transactions* in 1724. He had already attained some celebrity, for in that year he was elected to the Royal Society of London.

In 1724 the scale of Fahrenheit's thermometer for meteorological purposes began at 0° and ended at 95°. He describes his scale as depending on the determination of three points: the lowest was the 0° and was found by a mixture of ice, water, and sea salt; the next was the 32° point and was found by dipping the thermometer in a mixture of ice and water without the salt; the third was the 96° point to which alcohol expanded "if the thermometer be held in the mouth or armpit of a healthy person." The divisions were called *degrees*. When this scale was extended, the boiling point was found to be 212°. It has since been determined that the normal temperature of the human body is not 96°, but 98.4°, on Fahrenheit's scale.

It has also been supposed that Fahrenheit's division of the scale from freezing to boiling into 180° was in imitation of the division of a semicircle into 180 degrees of arc. The division of a circle into 360° is a survival of the sexagesimal system, and is convenient because 360 has a large number of divisors.

The Centigrade scale was introduced by Celsius, professor of astronomy in the university of Upsala, about 1742. He divided the interval between the freezing and the boiling point into 100 equal parts, but he placed the 0° at the boiling point and the 100° at the freezing point. The inversion of this scale, making the 0° the freezing point, was due to Strömer, a colleague of Celsius, eight years later. The sim-

plicity of Celsius's mode of division of the distance between the two points of reference has led to its general adoption in all countries for scientific purposes.

The scale of Réaumur, on which the freezing point is marked 0° and the boiling point 80°, has nothing to recommend it except that he avoided the misplaced zero of Fahrenheit. It is said that "he found that spirit of wine, mixed with one fifth water, expanded between the freezing and the boiling temperatures of water from 1000 to 1080 volumes; so he divided the intervening distance on the stem into 80 parts." This scale is still used for domestic purposes in Germany, but usually alongside the Centigrade scale.

The number of thermometric scales used in the eighteenth century was at least nineteen. Fortunately all but three of them have passed into ancient history, and the sooner the Centigrade becomes the sole survivor the better.

Each of the three scales is extended beyond the fixed points as far as desired. The divisions below 0° are read as minus and are marked with the negative sign. The initial letters F., C., and R. indicate the Fahrenheit, the Centigrade, and the Réaumur scales respectively. In this book Centigrade degrees will be understood unless another scale is indicated.

369. Conversion of Readings from One Scale to Another.—AB in Figure 243 is a thermometer with the three scales attached, P is the head of the thread of mercury, and F, C, and R are the readings on the three respective scales.

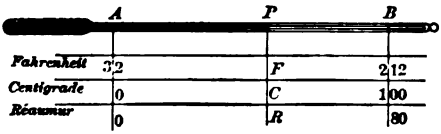


Fig. 243

On the Fahrenheit scale $AB = 180$ and $AP = F - 32$, since the zero is 32 divisions below A; on the Centigrade $AB = 100$ and $AP = C$; on the Réaumur $AB = 80$ and $AP = R$. Then the ratio of AP to AB is $\frac{F-32}{180} = \frac{C}{100} = \frac{R}{80}$. By substituting in this

equation the reading on any scale, the equivalent on either of the others may easily be found. For Fahrenheit readings 32 must be subtracted algebraically to find the number of degrees between the freezing point and the reading. Thus 50° F. is $50° - 32° = 18°$ above freezing; and $-10°$ F. is $-10° - 32° = -42°$, or 42° below freezing.

As an example of conversion from one scale to another, if it is desired to express 68° F. in degrees on the Centigrade scale, then $\frac{68 - 32}{180} = \frac{C}{100}$ and $C = 20$.

370. Limits of the Mercurial Thermometer. — Since mercury freezes at $-38.8°$, the mercurial thermometer cannot be used for temperatures below that point. The scale may be extended downward by alcohol to about $-110°$, and by pentane to $-200°$. But the dilatation of these liquids is not uniform; and thermometers filled by them must be calibrated by comparison with a standard.

Mercury boils under atmospheric pressure at about 350°. For higher temperatures up to 500°, mercurial thermometers may be used if the space above the mercury is filled with nitrogen under pressure to prevent boiling (§ 405). This is the upper limit for mercury in glass because the glass softens at higher temperatures. If the tube is made of fused quartz, the temperature of the mercury under sufficient pressure may be carried as high as 700°. The portion of the scale above 100° must be graduated by comparison with the normal hydrogen scale.


Fig. 244

371. The Clinical Thermometer. — The clinical thermometer in universal use among physicians and surgeons is a sensitive instrument of short range for indicating maximum temperatures. It is usually graduated from 95° to 110° F., or from about 35° to 45° C. There is a constriction in the tube just above the bulb (Fig. 244), which causes the thread of mercury to break at that point when the temperature begins to fall, leaving the top of the disengaged thread to mark the highest temperature registered. The

mercury can be forced down past the constriction by utilizing its inertia in tapping or jarring the thermometer.

Fig. 245

372. Beckmann's Thermometer. — To increase the length of a division for a Centigrade degree, so that it may be divided into hundredths on the scale, the bulb must be large and the capillary tube very narrow. The range of such an instrument is then quite limited, seldom exceeding ten degrees. To extend its availability Beckmann's thermometer has a reversed bulb at the top as a reservoir for excess mercury (Fig. 245). The bulb is 10 or 12 mm. in diameter and the stem is graduated to hundredths of a degree. The range is usually six degrees. When it is desired to use this thermometer for temperatures above 6°, it is raised to a temperature a little above the one to be measured, the excess mercury flowing into the upper bulb. Then the thread is detached with a shake, and the mercury remaining in the bulb and tube is used in the ordinary way. Such a thermometer is commonly employed for measuring small differences of temperature very accurately; the temperature corresponding to any point on the scale may be determined by comparison with a standard instrument.

373. Galileo's Air Thermometer. — The first air thermometer was invented by Galileo and it served in his time for the detection of fever. In its early form it consisted of a glass bulb at the end of a narrow tube, which was supported vertically in front of a scale and dipped into a vessel of colored liquid (Fig. 246). A small portion of the air in the bulb is expelled by warming, so that the colored alcohol will rise in the tube when the bulb again cools. If now the temperature rises, the liquid column falls; if the temperature falls, the liquid column rises. The instrument is remarkable for its sensitiveness, but until it has been greatly modified, it is only a *thermoscope*, because its readings change with barometric pressure as well as with temperature.

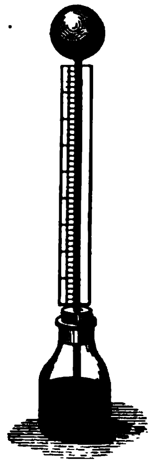


Fig. 246

The practical methods of using a gas as a thermometric substance are described in memoirs and large treatises. A description of a simple form of gas thermometer and the definition of the "absolute zero" are reserved for a later section.

III. EXPANSION

374. Expansion of Solids. — It is a familiar fact that solids in general expand when heated and contract again when their temperature falls. Gravesande's ring (Fig. 247) illustrates the expansion of a metal ball which just passes through the ring when both are at room temperature; but if the ball is heated in boiling water, it will no longer pass through the ring until it has again cooled.

Fig. 247

If a strip of sheet iron and one of copper are riveted together (Fig. 248) and supported at the ends, heating with a spirit lamp or a Bunsen burner will cause the compound bar to bend into an arc of a circle with the copper on the convex side, because the expansion of the copper is greater than that of the iron.

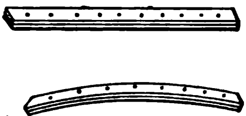


Fig. 248

A metal wire increases in length when heated. In Figure 249 an iron wire about a meter long is attached to a wooden support at one end A and at the other end to a screw eye in a long wooden pointer BC. The pointer is free to turn around a pin at B very near the screw eye, and its weight keeps the wire stretched. When the wire is heated by passing through it a suitable electric current, its expansion is indicated by a long sweep of the pointer. The wire cools quickly after the current is cut off and the pointer returns to its former position.

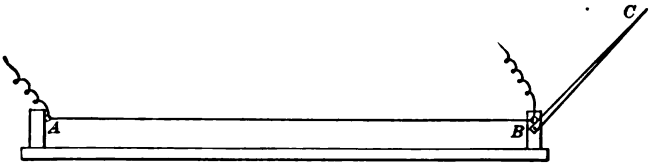


Fig. 249

Wagon tires are shrunk on wooden wheels by putting them on while red-hot and suddenly cooling. Steel rims are shrunk on street car wheels in the same manner. Large guns at the Washington Navy Yard are made

by slipping a red-hot outer cylinder over a cold inner one in a vertical position; when the hot cylinder cools, it contracts powerfully on the cold one.

375. Coefficient of Expansion.—The question as to how much a body changes in dimensions with a given change of temperature can always be answered if its coefficients of expansion are known.

Let d_0 be the numerical measure of any dimension of a body at 0°, d its value at $t°$, and Δd the change in this dimension per degree. Then the relation between the values of the dimension at the two temperatures may be expressed with sufficient accuracy for most purposes by the linear equation

$$d = d_0 + \Delta d t = d_0\left(1 + \frac{\Delta d}{d_0} t\right).$$

If now θ be substituted for the ratio $\Delta d/d_0$, the equation becomes

$$d = d_0(1 + \theta t). \tag{69}$$

The *coefficient of expansion* θ may be defined as *the ratio of the increase in the dimension per degree to the dimension at zero.*

376. Coefficients of Expansion of a Solid. — A solid has three coefficients of expansion, according as d is regarded as a length, an area, or a volume. Equation (69) has three corresponding forms,

$$\left.\begin{aligned} l &= l_0(1 + \lambda t) \\ s &= s_0(1 + \sigma t) \\ v &= v_0(1 + \alpha t) \end{aligned}\right\}. \tag{70}$$

The constants λ, σ, α are the coefficients of linear, superficial, and cubical expansion respectively.

The relation between these three coefficients is a simple one in the case of isotropic bodies which expand equally in all directions, that is, have the same linear coefficient in three directions at right angles. Then the coefficient of expansion in volume is three times that in length. The volume of a

cube whose edge is l_0 at 0° is $l_0^3(1+\lambda t)^3$ at t°. It is also $v_0(1+\alpha t)$. Equating these values, and remembering that $v_0 = l_0^3$, we have

$$1+\alpha t = (1+\lambda t)^3 = 1 + 3\lambda t + 3\lambda^2 t^2 + \lambda^3 t^3.$$

But since λ is a number of the order of 1/100,000, its higher powers may be neglected in comparison with the first, or

$$1+\alpha t = 1 + 3\lambda t, \text{ and } \alpha = 3\lambda \text{ nearly.}$$

In a similar manner it may be shown that $\sigma = 2\lambda$.

In general crystals have three axes of expansion at right angles, and the three linear coefficients in these directions are not identical. The coefficient of volume is then equal to the sum of the three linear coefficients. In such cases a crystalline sphere at one temperature ceases to be spherical at other temperatures; also a crystalline body in the form of a cube at one temperature does not remain cubical when the temperature changes, unless the crystal belongs to the cubic system.

When quartz crystals are strongly heated, their unequal expansion in different directions causes them to burst into small pieces. Large diamonds have been known to break into fragments by the heat of the hand when they are first taken from the earth.

LINEAR COEFFICIENTS OF SOLIDS

(Centigrade Scale)

Glass	0.0000088	Mercury (cubical)	0.000180
Platinum	0.0000089	Quartz ∥ axis	0.000007
Copper	0.000017	Quartz ⊥ axis	0.000014
Gold	0.000015	Quartz fused	0.0000004
Iron	0.000012	Ebonite	0.00008
Lead	0.000029	Porcelain	0.0000027
Silver	0.000019	Nickel steel (36% Ni)	0.0000009
Zinc	0.000030		

377. Applications of Expansion. — In addition to the instances of expansion already mentioned, the following more or less familiar facts

may be accounted for by the expansion taking place with changes of temperature:

When hot water is poured into a thick glass vessel, especially if it is not well annealed and so under stress, as shown by polarized light (§ 348), it will probably break, because the temperature difference between the outer and inner portions produces stresses greater than the glass can sustain. The low temperature of a mixture of solid carbon dioxide and sulphuric ether is equally destructive.

After quartz has been fused it loses its crystalline structure and its linear coefficients of expansion become equal and very small. It is for this reason that vessels made of fused quartz may be raised to a red heat and plunged into cold water without breaking.

Fig. 250

The inequality in the coefficients of expansion of different substances is utilized in thermometers, in the compensated clock pendulum, and in the compensated balance wheel of a watch.

Graham's mercurial pendulum bob consists of one or more glass jars, nearly filled with mercury, and attached to the lower end of the pendulum rod (Fig. 250). A rise of temperature lengthens the rod, and lowers the center of oscillation of the pendulum; but the mercury expands upwards and compensates by raising the center of oscillation. The adjustment of the compensation is made by raising or lowering the cylinder of mercury by means of a screw.

The rate of a watch or a chronometer depends largely on the balance wheel. Unless this is compensated, it expands when the temperature rises and the watch loses time, the larger wheel oscillating more slowly under the force supplied by the elasticity of the hairspring. Compensation is effected by making the rim of the wheel in two or more sections, each being made of two materials (like Fig. 248) and supported by one end on a separate arm (Fig. 251). The more expansible metal is on the outside. Then when the temperature rises and the wheel as a whole expands, the loaded ends a, a' of the sections move inward, thus compensating for the increase in the length of the radial arms. The small screws on the rim are for the adjustment of the moment of inertia of the wheel. In practice the compensation is seldom perfect, and the rate of the chronometer must be checked by stellar observations, that is, by means of the uniform rotation of the earth on its axis.

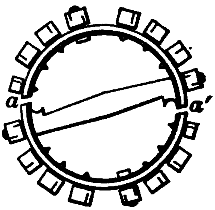


Fig. 251

It has been found possible to make a nearly perfect compensation for expansion by a particular percentage composition of an alloy of steel and

nickel, named "invar" by its inventor, Guillaume of Paris. It contains 36 per cent of nickel and its coefficient of linear expansion is only one tenth that of glass and less than one part in a million. This property commends it for use in the construction of standards of length, of pendulums, and for other instruments in which invariability is highly desirable.

378. Expansion of Liquids.—In the case of liquids and gases expansion in volume only comes into consideration. With the exception of water and aqueous solutions, the expansion of fluids is positive and larger than that of solids. Two values of this coefficient for fluids must be carefully distinguished, the *absolute* and the *apparent*. The former is defined by the third equation of (70); the latter is the observed coefficient when its value is affected by the expansion of the containing vessel. The absolute coefficient is the sum of the apparent coefficient and the coefficient of the vessel.

The expansion of liquids in general increases with the temperature and becomes large at high temperatures. The volume of liquids cannot be expressed with one coefficient as a linear function of the temperature, but it requires two or more constants in the equation connecting volume and temperature. The equation has the form $v = v_0(1 + \alpha t + \beta t^2 + \gamma t^3 + \cdots)$. The coefficients above the first are usually relatively small, but they are not negligible.

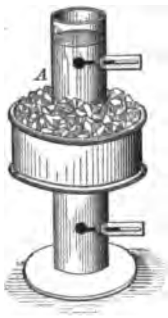


Fig. 252

379. Expansion of Water.— Water exhibits the remarkable property of contracting when heated at the freezing point. This contraction continues up to 4°; at this temperature expansion sets in, so that the greatest density of water is at a temperature of 4°, and its density at 6° is nearly the same as at 2°.

The peculiar behavior of water is illustrated by Hope's apparatus (Fig. 252). It consists of a glass jar with thermometers inserted near the top and bottom. Around the middle is an annular reservoir. If the vessel is

filled with water at about 10°, the upper thermometer will show at first a slightly higher temperature than the lower one. If now the trough at the middle be filled with a freezing mixture, the first effect will be the gradual fall of the lower thermometer to 4° without much change of the upper one. After the lower thermometer becomes stationary, the upper one falls rapidly until its reading is reduced to zero and ice forms at the surface. The water at 4° sinks to the bottom, while that below 4° is lighter and rises to the top, where the freezing begins. For this reason ice forms at the surface of a body of cold water, which freezes from the surface downward, instead of from the bottom upward.

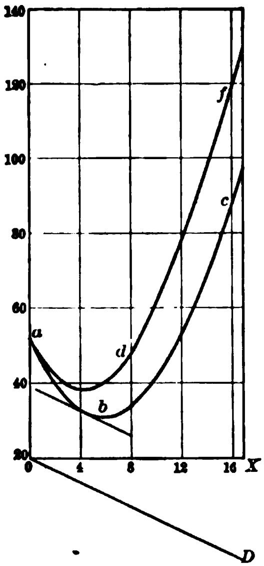


Fig. 253

The relation between the volume and the temperature of water near the freezing point may be determined by means of a large thermometer filled with distilled water. If the apparent volumes of the water in glass are plotted as ordinates and the corresponding temperatures as abscissas, the curve is approximately a parabola *abc* (Fig. 253). The vertex is somewhat above 4° and this is the temperature of the least *apparent* volume. The observations for this curve include the dilatation of both the glass and the water. The true volume-temperature curve for water may be found by adding to the ordinates of this curve the expansion of the glass. For this purpose, it is only necessary to draw a line *OD*, making with the axis of temperatures an angle whose tangent, expressed in terms of the vertical and horizontal scales, is the dilatation of glass for one degree. If then the vertical distance between *OX* and *OD* are added to the corresponding ordinates of *abc*, the result

is the curve *adf*. The point of least volume and greatest density corresponds to the shortest vertical distance between *OD* and the curve *abc*. This may be found by drawing a tangent to the curve parallel to *OD*. The tangent touches the curve very near the ordinate passing through 4°, and this is the temperature of least volume.

380. Expansion of Gases — Law of Charles. — The law of expansion of gases, discovered by Charles in 1787 and confirmed by Gay-Lussac in 1802, is the following: The volume of a given mass of any gas, under constant pressure, increases from the freezing to the boiling point by a constant fraction of its volume at zero. This law is known as the law of Charles or of Gay-Lussac. For the Centigrade scale the constant fraction is 0.3665 for dry air. This is equivalent to 0.003665 for one degree C. A near approximation is 1/273. Therefore 30 cm.^{3} of a gas at 0° become about 41 cm.^{3} at 100°.

It follows from the law of Charles that the expression for the dilatation of a solid may also be applied to that of a gas under constant pressure, or

$$v = v_0 (1 + \alpha t) = v_0(1 + 0.003665\, t). \qquad (71)$$

The investigations of Regnault and others have shown that this law, like that of Boyle, is not exact, but is a close approximation only in the case of gases which most nearly obey Boyle's law.

A gas which would obey Boyle's law exactly is known as an ideal or *perfect gas*. For such a gas, the product pv of the pressure and the volume, for a constant temperature, is a constant. It follows that this product is some function of the temperature, or

$$pv = f(t).$$

It is obvious from this expression that the changes produced by the application of heat to a gas may be investigated either by observing the changes of volume under constant

pressure, or the changes of pressure at constant volume. The two methods have been found to give nearly identical results.

COEFFICIENTS OF EXPANSION AND PRESSURE

	Constant Pressure	Constant Volume
Hydrogen	0.003660	0.003663
Helium	3663	3664
Air	3671	3665
Nitrogen	3671	3668
Carbon dioxide	3710	3686
Cyanogen	3877	3682
Sulphur dioxide	3903	3670

The easily liquefiable gases at the bottom of the list have a larger coefficient of dilatation than those which are liquefied with great difficulty. It was found that the coefficient of the gases investigated approach equality as the initial pressure decreases; and the more highly rarefied they are, the more nearly do they approach the ideal limit of exact obedience to Boyle's law.

381. The Absolute Scale of Temperature. — The law of Charles leads to another scale of temperature known as the *absolute scale.* By this law the volumes of any mass of gas, under constant pressure at 0°, and at any other temperature $t°$, are connected by the relation

$$v_t = v_0 + \text{the expansion} = v_0\left(1 + \frac{t}{273}\right).$$

Also $$v_{t'} = v_0 + \text{the expansion} = v_0\left(1 + \frac{t'}{273}\right).$$

Then $$\frac{v_t}{v_{t'}} = \frac{v_0\left(1 + \frac{t}{273}\right)}{v_0\left(1 + \frac{t'}{273}\right)} = \frac{273 + t}{273 + t'}$$

Suppose now a new scale be chosen, whose zero is 273 degrees Centigrade below the freezing point of water, and let

temperatures on this new scale be denoted by T. Then $273 + t$ will be represented by T, and $273 + t'$ by T' on the new scale, and

$$\frac{v_t}{v_{t'}} = \frac{273 + t}{273 + t'} = \frac{T}{T'},$$

or, *the volumes of the same mass of gas, under constant pressure, are proportional to the temperatures on this new scale.* The point 273° below 0° is called the *absolute zero*, and the temperatures on this scale, *absolute temperatures.* Up to the present time it has not been found possible to cool a body to the absolute zero; but by evaporating liquid hydrogen under low pressure, Sir James Dewar reached a temperature which he estimated to be within 9° of the absolute zero. In physical theory the absolute zero is more important, and, for practical purposes, it is more convenient, than the arbitrary zero of the Centigrade scale.

382. The Laws of Boyle and Charles Combined. — The laws of Boyle and Charles may be combined into one expression, which is known as *the gas law*, though it has greater generality than its method of derivation from gases would imply.

Let v_0, p_0, T_0 be the volume, pressure, and absolute temperature of a gas under standard conditions of 0° temperature and 76 cm. of mercury pressure.

Also let v, p, T be the corresponding quantities at temperature T.

Then, applying Boyle's law to increase the pressure to the value p, the temperature remaining constant, we have

$$v_0 : v' :: p : p_0.$$

By changing the pressure from p_0 to p, the volume has become v'.

Next apply the law of Charles, keeping the pressure constant at the value p, and starting with the value v'. Then

$$v' : v :: T_0 : T.$$

These changes have taken place by two successive steps. From the first proportion $\frac{v_0}{v'} = \frac{p}{p_0}$; and from the second $\frac{v'}{v} = \frac{T_0}{T}$. Multiplying the two equations together member by member, and

$$\frac{v_0}{v} = \frac{pT_0}{p_0T}, \text{ or } \frac{p_0v_0}{T_0} = \frac{pv}{T} = \text{a constant.}$$

This constant is usually denoted by R. We may therefore write

$$pv = RT. \tag{72}$$

Since R in this equation is a constant, it follows that in a perfect gas obeying the laws of Boyle and Charles, both the pressure at constant volume and the volume under constant pressure are directly proportional to the absolute temperature.

383. **Numerical Value of the Constant** R. — If v in equation (72) is the volume containing the unit mass of the gas under standard conditions of temperature and pressure, the constant R will be inversely as the density of the gas; for, with p and T constant, R is proportional to v, and v is inversely as the density d.

If, however, the same volume v be taken for different gases, then R will have the same value for all. But what shall this volume be? The hypothesis of Avogadro is that equal volumes of all gases at the same temperature and pressure contain the same number of molecules. The masses of equal volumes are therefore proportional to the molecular weights. The particular volume chosen for the gas equation is the volume containing the *gram molecule*, that is, a number of grams equal to the molecular weight.

The constant value of R may be found from the following data:

The standard pressure p of one atmosphere is 1,013,250 dynes per square centimeter (§ 171). The density of oxygen at 0° and under a pressure of one atmosphere is 0.0014279 gm. per cubic centimeter, and its molecular weight is 32.

Hence the volume v containing 32 gm. is $32/0.0014279 = 22{,}410$ cm.³. Therefore

$$R = \frac{pv}{T} = \frac{1013250}{273} \times 22{,}410 = 83{,}175{,}000. \qquad (73)$$

Since T is a number and the product pv is work,

$$R = 8.317 \times 10^7 \text{ ergs} = 8.317 \text{ joules/degree.}$$

384. The Constant Volume Gas Thermometer.—In the gas thermometer the increase in temperature is assumed to be proportional to the increase in pressure of the gas when its volume is kept constant.

The dry gas is inclosed in a suitable bulb a (Fig. 254), which is connected by means of a capillary tube to one of larger cross section at B. The tubes CE and BD are joined by means of stout rubber tubing, which permits CE to be raised or lowered so as to keep the mercury surface at B. In the larger tube at B is a hook pointing downward; the surface of the mercury is always brought just in contact with this hook before a reading is taken. The volume of the gas is then a constant, subject to a small correction for the expansion of the glass. The difference in the level of the mercury at B and E, added to the height of the barometer, both corrected for temperature, gives the pressure under which the gas is in the bulb.

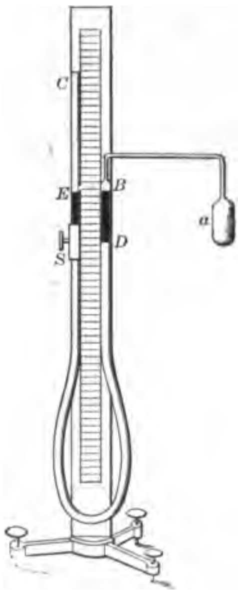


Fig. 254

Since hydrogen at high temperatures diffuses through the walls of a containing vessel, nitrogen is sometimes used instead. For temperatures between 0° and 100° the difference between the readings on the normal hydrogen scale and those with nitrogen are at most only a

few hundredths of a degree; at both lower and higher temperatures they become appreciable.

If p_0 is the pressure at 0° and p the pressure at some higher temperature $t°$, then since the absolute zero is 273° below the zero of the Centigrade scale, we may write

$$\frac{273 + t}{273} = \frac{p}{p_0}. \quad \text{Whence } t = 273\left(\frac{p}{p_0} - 1\right). \tag{74}$$

The pressure at zero is determined by surrounding the bulb of the thermometer with melting ice and taking readings. Any other temperature is then measured by observing the pressure necessary to keep the surface of the mercury at the fiducial point near B.

IV. MEASUREMENT OF HEAT

385. Unit Quantity of Heat.—For the measurement of heat as a physical quantity no knowledge of the ultimate nature of heat is required; the methods of measurement are based on some property or effect attributed to heat. The measurement of heat is called *calorimetry*.

Heat, like other physical quantities, must be expressed in terms of a convenient unit. The unit quantity of heat is the heat required to raise the temperature of unit mass of water one degree. If the unit mass is the gram and the unit of temperature the degree Centigrade, the unit of heat is called the *calorie*. In engineering practice in England and America the British thermal unit (B. T. U.) is commonly employed. It is the heat required to raise the temperature of one pound of water one degree Fahrenheit.

Since the quantity of heat which will warm a gram of water one degree at different temperatures between 0° and 100° is not rigorously the same, the temperature on the scale should be stated for an exact definition of the calorie. No agreement with respect to this point has been reached by physicists. If the interval is chosen from 15° to 16°, the calorie will then be the one hundredth part of the heat re-

quired to raise the temperature of one gram of water from 0° to 100°.

386. Thermal Capacity. — The *thermal capacity* of a *body* is the number of heat units required to raise its temperature one degree. The thermal capacity in calories of any mass of water is numerically equal to that mass in grams.

The thermal capacities of equal masses of different substances differ widely. Thus, if 100 gm. of mercury at 80° be mixed with 100 gm. of water at 20°, the temperature of the whole mass will be about 22°. The heat given up by the mercury in cooling 58° heats the same mass of water only about 2°. The thermal capacity of the water is about thirty times as great as that of the mercury. So also if 100 gm. of copper at 100° be cooled in 100 gm. of water at 0°, the final temperature will be only 9.1°. The heat lost by the copper in cooling through 90.9° is sufficient to raise the temperature of the same mass of water only 9.1°.

387. Specific Heat. — The *specific heat* of a *substance* is the ratio between the heat capacities of equal masses of the substance and of water. This amounts to saying that the thermal capacity of unit mass of a substance is its specific heat.

For example, the heat capacity of one pound of iron is 0.112 B.T.U. Also, the heat capacity of one gm. of iron is 0.112 calories. Its specific heat is therefore 0.112. Specific heat is a ratio and is independent of the unit of heat measurement. The thermal capacity of a body is the product of its specific heat and its mass. The numerical value of thermal capacity depends on the units employed.

388. Specific Heat by the Method of Mixtures. — The method most commonly applied to determine specific heats is the method of mixtures. It is based on the experimental fact that when an exchange of heat takes place between bodies in thermal contact, the quantity of heat lost by one part of the system is the same as that gained by the other. In apply-

ing this principle of equal heat exchanges, it is necessary to guard against its application in cases where there is either absorption or generation of heat, due to some reaction taking place between the parts of the mixture. For example, it cannot be applied to the determination of the specific heat of a salt solution by mixing it with water, for in general there is a change in temperature, due to the heat of dilution, when a salt solution is mixed with water, both at the same temperature.

Let two bodies A_1 and A_2 have masses m_1 and m_2, temperatures t_1 and t_2, and specific heats s_1 and s_2. If they are placed in thermal contact, they will reach some intermediate temperature t, which is higher than t_1 and lower than t_2. Then the quantity of heat lost by A_2 will be $m_2s_2(t_2-t)$, and the heat gained by A_1 will be $m_1s_1(t-t_1)$. Assuming now no generation or absorption of heat and that the only exchange is between A_1 and A_2, the heat lost by one is gained by the other, or

$$m_2s_2(t_2-t)=m_1s_1(t-t_1) \text{ and } \frac{s_2}{s_1}=\frac{m_1(t-t_1)}{m_2(t_2-t)}.$$

If A_1 is water and its specific heat is unity, the last equation becomes $s_2=\dfrac{m_1(t-t_1)}{m_2(t_2-t)}$.

In practice it is necessary to take into account the thermal capacity of the calorimeter itself, since both the vessel and its contents change temperature equally. The number of heat units required to raise the temperature of the calorimeter one degree is called its *water equivalent*, because it is equal to the mass of water having the same thermal capacity as the calorimeter.

Let the water equivalent be m. Then the heat gained by the calorimeter and its contents is $(m+m_1)(t-t_1)$, and the specific heat s_2 becomes

$$s_2=\frac{m+m_1}{m_2}\cdot\frac{t-t_1}{t_2-t}. \qquad (75)$$

For the radiation correction for the exchange of heat between the calorimeter and its surroundings the reader is referred to laboratory manuals.

SPECIFIC HEATS AT 50°

Antimony	0.051	Mercury	0.033
Bismuth	0.031	Platinum	0.032
Cadmium	0.059	Silver	0.056
Copper	0.095	Tin	0.054
Glass (Jena 16^{III})	0.199	Zinc	0.093
Iron	0.112	Ice (−30°—0°)	0.505
Lead	0.031	Alcohol (10°—15°)	0.602
Magnesium	0.245	Turpentine	0.468

389. Specific Heat of Water. — Water has a higher thermal capacity than any other substance except hydrogen. The specific heat of water is nearly twice as great as that of ice (0.505), and more than twice as great as that of steam under constant pressure (0.477).

The distribution of large quantities of heat in buildings by means of hot water is practicable because of its large thermal capacity. As it is, the radiating surface must be larger for heating by hot water than by steam.

The beneficent influence of the water of the ocean in equalizing climatic differences between summer and winter at once suggests itself. The ocean stores the heat of summer and gives it out gradually in winter. Hence the absence of extremes in an island climate.

The specific heat of water is not a constant. The exhaustive experiments of Rowland on the dynamical equivalent of heat were the first to demonstrate that the specific heat of water decreases from 0° to about 30° and then increases again. The precise point of this minimum value is difficult to determine, since the change in the specific heat near the minimum is very small.

390. Two Specific Heats of a Gas. — The specific heat of a gas is an indeterminate quantity unless the conditions are defined.

It may be measured in two ways: under the condition of a constant pressure, or at a constant volume. The former is the *specific heat under constant pressure*, and the latter the *specific heat at constant volume.* If the pressure is maintained constant, the volume must increase, and in consequence the gas does work (§ 66). The heat applied to a gas expanding under pressure not only raises the temperature and increases the kinetic energy of the molecules, but it also does the work of expanding under pressure. The specific heat under constant pressure is therefore greater than the specific heat at constant volume. The ratio between them is involved in the calculation of the velocity of sound in gases (§ 208).

The following table gives the value of the specific heat under constant pressure for several gases:

Air	0.2374	Chlorine	0.1240
Bromine	0.0555	Hydrogen	3.4090
Carbon dioxide	0.2169	Nitrogen	0.2438
Carbon monoxide	0.2405	Oxygen	0.2175

391. Thermochemical Equations. — A chemical reaction is nearly always attended by a thermal change, which is generally an evolution of heat. In equations denoting thermochemical reactions the symbols serve to indicate the substances taking part in the reaction as in ordinary chemical equations; they denote also definite quantities of these substances in grams equal to their atomic or molecular weight. Further, they are energy equations and are incomplete unless the energy involved in the reaction is also included in them. Thus

$$\text{Na} + \text{Cl} = \text{NaCl} + 20{,}400 \text{ cal.}$$

means that the union of 23.05 gm. of sodium and 35.45 gm. of chlorine produces 58.5 gm. of sodium chloride (one gram molecule), together with the generation of 20,400 calories of heat or the equivalent energy in some other form, such as that of an electric current. The 20,400 calories represent the difference between the energy associated with these definite quantities of sodium and chlorine on the one side, and of the sodium chloride on the other.

Thermochemical processes are called *exothermic* when heat is evolved, and *endothermic* when it is absorbed. In the former the energy in calories (or in joules) has the positive sign; in the latter, the negative.

The heat of reaction has different names according to the process producing it. Thus we have heat of formation, heat of hydration, heat of solution, heat of dilution, etc. The following examples illustrate them:

(1) Heat of formation:

$$Pb + 2\,I = PbI_2 + 39{,}800 \text{ cal. (exothermic).}$$
$$C + 2\,S = CS_2 - 28{,}700 \text{ cal. (endothermic).}$$

(2) Heat of hydration:

$$ZnSO_4 + 7\,H_2O = ZnSO_4 \cdot 7\,H_2O + 22{,}690 \text{ cal.}$$
$$CuSO_4 + 5\,H_2O = CuSO_4 \cdot 5\,H_2O + 18{,}550 \text{ cal.}$$

(3) Heat of solution and dilution:

$$ZnSO_4 \cdot 7\,H_2O + 393\,H_2O = ZnSO_4 \cdot 400\,H_2O - 4260 \text{ cal.}$$
$$CdSO_4 \cdot \tfrac{8}{3}\,H_2O + 397\tfrac{1}{3}\,H_2O = CdSO_4 \cdot 400\,H_2O + 2660 \text{ cal.}$$

(4) Heat of dilution:

$$ZnSO_4 \cdot 20\,H_2O + 380\,H_2O = ZnSO_4 \cdot 400\,H_2O + 400 \text{ cal.}$$
$$ZnCl_2 \cdot 5\,H_2O + 395\,H_2O = ZnCl_2 \cdot 400\,H_2O + 8020 \text{ cal.}$$

(5) Heat of combustion:

$$2\,H + O = H_2O + 69{,}400 \text{ cal.}$$
$$C + 2\,O = CO_2 + 103{,}000 \text{ cal.}$$

V. CHANGE OF STATE

392. Changes of State Produced by Heat. — When a crystalline solid is heated its temperature rises until it begins to pass into the liquid state. So long as the supply of heat is moderate, the temperature remains constant until the entire mass has *fused* or *melted*, after which the temperature rises again to the point at which boiling sets in. Again the temperature remains constant until the whole of the liquid has changed to vapor. The temperature will then again rise. At either stationary temperature and under a constant pressure the two states of the substance are in equilibrium with each other, and a mixture of the two will remain unchanged if no heat is supplied or lost.

Most substances are capable of existing in more than one state or modification. Thus, water may exist as ice, liquid water, or water vapor; and sulphur as a vapor, a liquid, or as either of two solids crystallizing in different systems. Such modifications, when they exist together, and may be separated from each other, as ice from water, are known as *phases*. The physical conditions affecting equilibrium of these phases are temperature and pressure.

At 0° and under a pressure of one atmosphere ice and water have the same vapor pressure and are in equilibrium. They may exist together at this temperature and pressure in all relative proportions, and these proportions will remain unchanged so long as the mixture neither gains nor loses heat. So also at 100° and under atmospheric pressure, water and water vapor exist together, and the two phases are in equilibrium.

393. The Melting Point. — The *melting* or *fusing* point of a substance *is the temperature at which the solid and liquid states or phases are in equilibrium* under some defined pressure, usually atmospheric. Above this temperature the substance is a liquid; below it, normally a solid. Solidification or freezing is the converse of fusion, and the temperature of solidification of any substance is normally the same as the melting point.

The melting point of ice is sharply marked, and there is no appreciable difference of temperature between the melting ice and the liquid phase into which it passes. This is generally true of crystalline substances, but the case is very different with amorphous solids, like wax, glass, and wrought iron, which cannot be said to have a definite melting point. Such substances soften and become plastic before reaching a more or less viscous liquid state. It is on account of this property that glass can be bent, moulded, drawn out into rods and tubes, or blown into various forms. Similarly the softening of wrought iron at a temperature much below the

melting point permits the metal to be rolled, forged, and welded. In the fusion of wax the outer portions are softer than the interior and are at a slightly higher temperature. The experiments of Person go to show that ice begins to increase in specific heat between $-2°$ and $0°$, and that there is a very small range of temperature within which it softens and melts. The difference between it and wax from this point of view is one of degree.

In general crystalline bodies have a definite fusing point, or a temperature at which they may exist either as a solid or a liquid; amorphous bodies, on the other hand, pass gradually from the solid to the liquid phase. A liquid which passes abruptly from the solid to the liquid state may be carefully cooled several degrees below the normal freezing point without solidifying. Thus water, protected from vibrations or covered with oil, will remain liquid at $-10°$, or in capillary tubes at $-20°$. This phenomenon is called *undercooling*.

Undercooled liquids are in unstable equilibrium; for if they be jarred, or if a solid fragment of the same substance be dropped into the liquid, solidification takes place rapidly with disengagement of heat, the temperature rising to the normal freezing point.

As examples, water in a small bulb containing also a thermometer may be cooled oy a freezing mixture to $-8°$ or $-10°$. When it finally freezes the temperature rises rapidly to $0°$.

Hyposulphite of sodium may be melted at $50°$ and then cooled without solidifying to room temperature. The addition of a small crystal of the salt destroys the equilibrium, solidification sets in around the crystal, and the temperature rises to the melting point of $47.9°$. So also sulphate of sodium, $NaSO_4 \cdot 10\,H_2O$, which has a constant melting point of $32.38°$, may be undercooled, and when it solidifies its temperature rises to the normal melting point, for this is the only temperature at which the solid and liquid can exist permanently in contact.

394. Freezing Point of Solutions. — The liquid solution of any substance has a lower freezing point than that of the pure solvent. An aqueous solution of any salt must be cooled

below the freezing point of water before ice separates out; for small concentrations the lowering of the freezing point is proportional to the amount of the substance in solution. The *freezing point of a solution* is the temperature at which the solution is in equilibrium with the solvent in the solid form.

Practical applications of the lowering of the freezing point by dissolving substances in water are familiar. Snow or ice is melted by salt, and alcohol or glycerine added to water lowers the freezing point. A so-called non-freezing mixture for the radiator of a motor car is made by adding alcohol to water. Twenty-five per cent of alcohol lowers the freezing point to $-13°$.

A dilute solution of common salt is in equilibrium with ice at a temperature a little below 0°. If now the temperature be reduced below this particular one for equilibrium, ice will separate, the concentration of the remainder of the solution will be increased, and the freezing point will be lower than that of the original solution. This process does not continue indefinitely because there is a limit to the solubility of salt in water, and this finite solubility sets a limit to the lowering of the freezing point. Beyond that limit, both ice and salt separate out and leave the concentration unchanged (§ 399).

395. Change of Volume in Fusion. — Most bodies occupy a larger volume in the liquid than in the solid state. Water, bismuth, and cast iron are exceptions. These expand when they solidify. The increase in volume when water freezes is very marked. Hence the bursting of water pipes, and the disruption of rocks by the freezing of water in cracks and crevices.

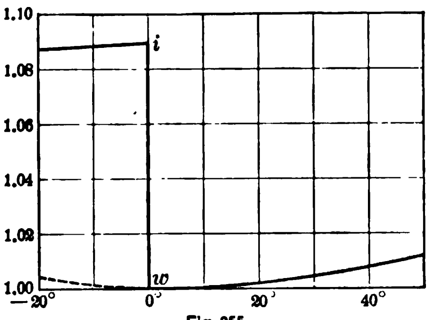


Fig. 255

The expansion of cast iron and type metal when they solidify contributes to an exact reproduction of the mould in which they are cast.

Figure 255 shows graphically the changes in the volume of

a gram of water from $-20°$ to $50°$. When water is undercooled, instead of the abrupt expansion at $0°$ accompanying freezing, the undercooled water expands gradually along the dotted curve, which has its vertex at $4°$.

396. Influence of Pressure on the Melting Point. — In 1849 Professor James Thomson showed from the dynamical theory of heat that a substance which expands on solidifying, like water, has its melting point lowered by pressure; but if the body contracts when it solidifies, like paraffin, then pressure raises its melting point. He calculated that in the case of water the freezing point is lowered $0.0075°$ by an increase of pressure of one atmosphere. Later Sir James Dewar found experimentally a reduction of $0.0072°$ per atmosphere up to 700 atmospheres. Hence, under a pressure of 1000 atmospheres water would not freeze above $-7.2°$. In other words, if water is confined so that it cannot freeze at $-7.2°$, it must be under a pressure of 1000 atmospheres.

A rough numerical statement is that a pressure of 145 kgm. per square centimeter (one ton per square inch) lowers the freezing point of water to $-1°$.

On the other hand, Bunsen found that the melting point of paraffin was raised from $46.3°$ to $49.9°$ by a pressure of 100 atmospheres.

397. Regelation. — Intimately connected with the lowering of the freezing point of ice by pressure is the phenomenon of re-freezing, or *regelation*, when the pressure is relieved. Regelation was first observed by Faraday.

Familiar examples are the hardening of snowballs under the pressure of the hands, and the formation of compact ice in a roadway, where it is compressed by vehicles and the hoofs of horses. Frozen foot-forms may often be seen to persist in compact ice after the loose snow has melted around them. The descending mass of snow from high altitudes becomes ice after melting under the pressure of its own weight and relief at lower levels. Such solidification occurs

only when the snow is soft and near its melting point. In cold weather snow will not pack. If two pieces of ice are firmly pressed together, even under warm water, they will adhere by regelation when the pressure is removed.

A glacier makes its way down its course by very irregular movements. Ice is probably plastic to some extent, but without doubt regelation plays an important rôle in glacial motion. The ice melts where it is subjected to enormous pressure by the descending masses above it. The melting permits the ice to accommodate itself to abrupt changes in the rocky channel, and a slow ice-flow results. As soon as the pressure at any surface is relieved, the water again freezes. The motion thus takes place by alternate melting and freezing. The middle of the flow moves faster than the borders of the stream, because the pressure there is greater and the consequent melting is more extensive.

Bottomley's experiment to illustrate regelation is very instructive. A stout block of ice is supported horizontally by wooden supports at its two ends. On it is hung a weight by means of a wire passed over the ice at the middle. The pressure melts the ice under the wire, and the water, passing around it and relieved of the stress, again freezes. In this way the wire cuts its way down through the ice, but the block of ice remains intact, though the track of the wire through it remains visible. The ice is as liable to break at any other place as along the track of the wire.

398. Heat of Fusion. — Since the temperature of a crystalline solid remains constant while heat is applied to melt it, it is obvious that heat disappears during the process. The heat that fuses a crystalline solid without raising its temperature becomes potential energy in doing the work of changing its state. When the liquid solidifies, this work is restored as heat.

The *heat of fusion* is defined as the number of heat units required to convert unit mass of the substance from the solid to the liquid state, without change of temperature. The whole quantity of heat H required to melt a mass m of any substance is proportional to its mass, or

$$H = lm. \tag{76}$$

The proportionality factor l is the heat of fusion. It is numerically equal to the number of units of heat necessary to melt unit mass of the substance. The unit in the *c.g.s.* system is the calorie per gram.

The heat of fusion may be determined by the method of mixtures. Let m_1 be the mass, and t_1 the temperature of the water and the calorimeter; also let m be the water equivalent of the calorimeter and m_2 the mass of the ice, the heat of fusion of which is to be found. If the temperature of the mixture after the ice is melted is t, then the heat lost by the calorimeter and its contents may be equated to the heat of fusion of the ice and its gain in heat in rising from 0° to $t°$,

or $$(m + m_1)(t_1 - t) = lm_2 + m_2t.$$

Whence $$l = \frac{(m + m_1)(t_1 - t)}{m_2} - t. \tag{77}$$

The most probable value for the heat of fusion of ice is 80 calories per gram, on the basis of a minimum specific heat of water at about 30°.

When the heat of fusion of ice is known, the ice calorimeter furnishes one of the standard methods for the determination of the specific heat of any substance which does not dissolve in water. The body of mass m_2 is placed in a cavity made in a block of ice of temperature 0°, and the mass m_1 of the ice that melts, while the body cools from temperature $t°$ to 0°, is measured. Then $$sm_2t = lm_1.$$

The specific heat s is given by this equation.

399. Heat lost in Solution. — Heat is absorbed when a body passes from the solid to the liquid state, although no heat is applied. If the liquefaction is brought about by solution in a proper solvent without chemical action, heat is still required to give mobility to the molecules, and the temperature of the solution falls. A thermometer will show a sensible fall of temperature when some finely divided ammonium nitrate is added

to a little water in a test glass. If a delicate thermoscope is used, such as a thermopile (§ 573), and a galvanometer, the heat absorbed by the solution of sugar in water may be detected. A still larger effect is produced by dissolving common salt, while quite a notable reduction of temperature is produced by dissolving nitrate of sodium in water.

A converse experiment, designed to show that heat is evolved when a substance becomes a solid by crystallization from a solution, is easily arranged by making a saturated solution of sodium hyposulphite at 30° in a small flask, and slowly cooling to about 20°. The solution is then undercooled. As soon as a very small crystal of the salt is dropped in, rapid crystallization sets in and extends through the whole solution. A thermometer in the solution shows that the temperature rises at the same time to about 30°.

Freezing mixtures are based on the absorption of heat necessary to give fluidity. Salt water freezes at a lower temperature than fresh water. When salt and snow or pounded ice are mixed together, both become fluid and absorb heat in the transition from one state to the other. By this mixture a temperature of $-22°$ may be obtained.

Figure 256 shows graphically why the temperature of a mixture of ice and salt water cannot be lower than $-22°$. The curve *AB* has freezing points as abscissas and concentrations as ordinates. The curve *CD* is the solubility curve of common salt (sodium chloride). The concentration of the salt solution decreases with falling temperature.

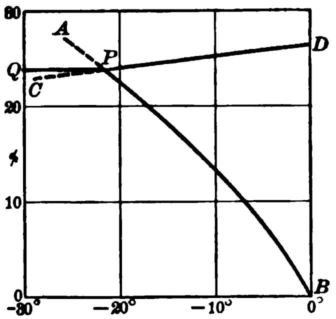


Fig. 256

From the point *B* to the intersection of the two curves, or from 0° to $-22°$, ice separates, leaving the solution increasingly concentrated with a continuous lowering of the freezing point. At *P* the solution has become saturated by the

removal of water as pure ice. After this point has been reached, ice and salt separate together in a constant proportion of about 25 per cent of salt to 75 of ice and at a constant temperature. The solid mixture is called a *cryohydrate*. After the solidification is complete, with continued cooling, the temperature falls along the line PQ; but then there is no longer a solution, but a mixture of ice and salt. Hence, $-22°$ is the lowest temperature at which a liquid salt solution can exist in equilibrium with ice. A freezing mixture should be composed of about one quarter by weight of salt and three quarters ice.

400. Vaporization. — *Vaporization* is the transition of a substance from the solid or the liquid state into that of a gas. There are four distinct forms of vaporization, depending on the conditions under which the transition occurs:

1. *Evaporation*, in which a liquid is converted into a vapor at its free surface at a relatively low temperature, and without the formation of bubbles.

2. *Ebullition*, or boiling, a rapid evaporation at a higher thermal equilibrium, when the liquid is visibly agitated by internal evaporation.

3. *Spheroidal state*, in which quiet evaporation, at a lower rate than boiling, goes on with the vapor acting as a cushion between the liquid and a surface at a relatively high temperature.

4. *Sublimation*, in which a solid passes directly into the gaseous form without passing through the intermediate liquid state.

Whether the gaseous condition is reached by one of these processes or another, heat is always absorbed, although the vapor formed is at the same temperature as the solid or the liquid from which it comes. The heat absorbed in the transition into the form of a gas is called the *heat of vaporization*.

401. Evaporation in a Closed Space. — The dynamical theory of heat assumes that the molecules of a liquid are in a state

of incessant agitation. The free path of the molecular motion in liquids is at least very limited; the migratory track of any individual molecule depends on its innumerable encounters with other molecules. While the average molecular velocity is determined by the temperature, the velocity of individual molecules may greatly exceed this average; and whenever a molecule at the free surface of a liquid has a normal component of velocity sufficient to carry it through the surface film, it escapes into the free space above the liquid. The concurrent escape of many molecules in this manner constitutes evaporation.

When evaporation takes place in a limited closed space, the free molecules may again return into the liquid. This return is called *condensation.* When the number of molecules making their escape equals the number returning through the surface film, the vapor in contact with the liquid is said to be *saturated.* It is in equilibrium with its liquid, and its pressure is the greatest it can have at the existing temperature. This maximum pressure is the *vapor pressure* at that temperature.

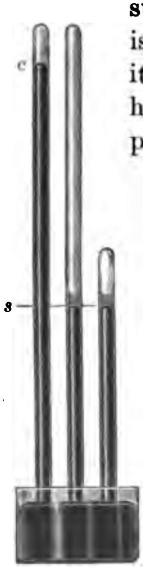


Fig. 257

The saturated vapor pressure is independent of the volume and depends only on the substance and its temperature. The value of this pressure and its independence of the volume are easily shown by means of Torricellian tubes. If the three tubes of Figure 257 are filled with mercury free from air, and are then inverted in a vessel of mercury, the mercurial columns in the long tubes will settle down to the height of the barometer at the time. If now by means of a small pipette, with its tip bent upward, a little sulphuric ether is allowed to enter two of the tubes, it will quickly evaporate and the mercury in them will fall to the same level through *s*. The column of mercury *cs* measures the vapor pressure of the ether. The mercury surfaces in the two tubes containing

the ether will stand at the same level, provided the two are at the same temperature with only a small quantity of liquid ether above the mercury. A rise in the temperature of the ether increases the vapor pressure and causes a further depression of the column of mercury.

If the volume of the saturated vapor is diminished without change of temperature, some of the vapor will condense to a liquid; if the volume is increased, more of the liquid will evaporate so as to maintain the same vapor pressure.

402. Ebullition and the Boiling Point. — Evaporation absorbs heat. If only a moderate amount of heat is supplied to a liquid, the evaporation is confined to its surface; with an increased supply the evaporation increases until the rate at which heat is supplied equals the rate of loss by evaporation. At a still higher temperature, this equilibrium of quiet evaporation can no longer be maintained, and bubbles of vapor form in the interior and at points of contact with the walls of the containing vessel. If the vapor pressure is insufficient to support these bubbles as they rise into the cooler liquid, they collapse with the familiar sound of "simmering." At a slightly higher temperature evaporation takes place into the bubbles themselves, and they rise with sufficient buoyancy to break through the surface film and escape. This process of rapid evaporation from the interior, as well as at the surface, is called *ebullition* or *boiling*.

The temperature at which this new equilibrium is established is called the *boiling point* of the liquid. It is constant for the same pressure. The normal boiling point assumes the standard atmospheric pressure of 76 cm. of mercury.

The boiling point of a liquid is the temperature at which it gives off vapor at a pressure equal to that sustained by the surface of the liquid. *The vapor pressure in boiling equals the pressure of the surrounding atmosphere.*

403. Superheating. — When the air has all been boiled out of a liquid and the containing vessel is clean, the tem-

perature may rise several degrees above the normal boiling point before ebullition begins. It then proceeds with almost explosive violence, and continues at this high rate until the temperature of the liquid falls to the point corresponding to the existing pressure. Air-free water may thus have an abnormally high initial boiling point. In clean glass vessels this condition gives rise to "bumping." It may be prevented by placing in the vessel insoluble bodies with rough surfaces, or by introducing some fresh liquid containing air.

While the temperature of the liquid may thus be above the normal boiling point, that of the saturated vapor is normal. Determinations of the boiling point are therefore made with the thermometer bulb enveloped by the saturated vapor just above the liquid.

A liquid may be superheated by keeping it out of contact with its gaseous phase, just as it may be undercooled by excluding the solid phase (§ 393). Drops of water 10 mm. in diameter have been kept liquid at 120° when suspended in a mixture of linseed and clove oils, and drops 1 mm. in diameter remained liquid up to 178°. They exploded when touched with a glass rod or by the side of the vessel.

404. Boiling Point of Solutions. — The vapor pressure of a salt solution is lower and the boiling point higher than those of the pure solvent. From solutions which do not contain a volatile constituent, such as alcohol, only the solvent evaporates. When the solution becomes saturated, the excess of salt crystallizes out.

The temperature of an aqueous solution of common salt may be as high as 110°; the boiling point of a saturated solution of calcium chloride is 180°, but the steam from the boiling solution falls quickly to 100° under normal atmospheric pressure. In other words, the temperature of the saturated vapor is the boiling point of the solvent and not that of the solution.

405. Influence of Pressure on the Boiling Point. — The vapor pressure and the boiling point of a liquid rise together, the pressure rising more rapidly than the boiling point. The vapor pressure curve for water is plotted in Figure 258 with pressures in centimeters of mercury as ordinates and temperatures as abscissas. This curve divides the diagram into two regions: below it the conditions permit the existence of water as an unsaturated vapor only; above it, as a liquid only; along the curve the liquid and the saturated vapor exist together in equilibrium.

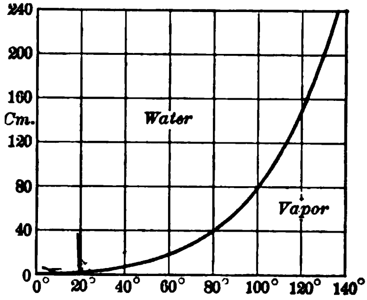


Fig. 258

At 9.2 cm. pressure the boiling point falls to 50° and at a pressure of 4.6 mm. it is the same as the freezing point. Near 100° the change in the boiling point is 0.1° for a change of pressure of 2.71 mm. of mercury. Hence at high altitudes water boils below 100°. At Quito the boiling point is about 90°.

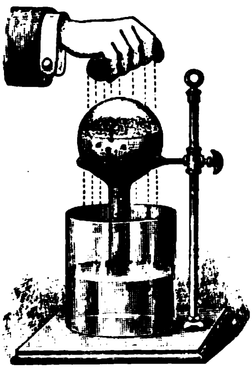


Fig. 259

Boiling under reduced pressure is resorted to in concentrating sugar solutions. The same principle is utilized in removing moisture from coils of insulated wire by heating them by means of steam pipes in an exhausted chamber.

The boiling of water at a reduced temperature may be conveniently shown by first boiling it in a round-bottomed flask, then corking tightly, inverting, and supporting in a ring stand (Fig. 259). The boiling may be renewed by applying cold water, which condenses the water vapor and reduces the pressure within the flask.

406. The Spheroidal State.—When a drop of water is carefully placed on a clean hot stove, it will often take a flattened globular form and roll about with rapid but silent evaporation. It is then in the *spheroidal state.* So liquid oxygen assumes the spheroidal state on water. It boils at $-180°$, and the water is at a high temperature relative to it. Spheroidal sulphur dioxide has a temperature low enough to freeze a drop of water placed in it. This may happen in a red-hot crucible because the sulphur dioxide in the spheroidal state is below its boiling point, and this is below the freezing point of water. The globular form of a spheroidal drop is accounted for by surface tension.

A liquid in a spheroidal state is not in contact with the hot surface, but rests on a cushion of its own vapor. If the drop is not too large, light may be projected through between it and a hot surface of platinum. The temperature of spheroidal water is from 90° to 98°. Under the receiver of an air-pump it may be as low as 80°.

407. Sublimation.—A body *sublimes* when it passes directly from the solid to the gaseous form without going through the liquid state. Ice and snow below freezing temperature gradually waste away by evaporation. Solid carbon dioxide disappears by sublimation without melting. It evaporates only as it gets heat to convert it into the gaseous form.

Camphor and ammonium carbonate sublime at room temperature. Iodine, ammonium chloride, and arsenic sublime when heated gently at atmospheric pressure. Arsenic may be fused if the pressure is increased.

Below a certain critical pressure for each, ice, mercuric chloride, and camphor cannot be melted, but they pass directly into the gaseous form. Whenever the boiling point of a substance may be lowered by reduced pressure to its freezing point, then at or below that pressure the solid will sublime. At a pressure less than 4.6 mm. of mercury ice is converted into water vapor by heat without melting.

408. Heat of Vaporization. — The *heat of vaporization* of a substance is the quantity of heat required to vaporize unit mass of it *without change of temperature*. The heat of vaporization of a liquid is usually understood to refer to its evaporation at its normal boiling point. In the *c.g.s.* system it is expressed in calories per gram.

The heat of vaporization varies with the temperature at which the vaporization takes place. The following expression is derived from the investigations of Regnault and Griffiths, and it applies to the evaporation of water between 0° and 100° C.:

$$L = 596.73 - 0.601\,t \text{ cal./gm. (calories at } 15°\text{)}.$$

L is the number of calories required to evaporate 1 gm. of water at $t°$. At 100°, therefore, the heat of vaporization is 536.6 cal./gm. To convert 1 gm. of water at 100° into steam at 100° requires 536.6 calories; conversely, when 1 gm. of steam at 100° condenses to water at the same temperature, 536.6 calories of heat are produced. The energy carried by steam is large because of the high heat of vaporization of water.

409. Cooling by Evaporation. — If the heat required to vaporize a liquid is not supplied from an external source, the evaporation will be accompanied by a lowering of its temperature. A little ether, alcohol, or gasoline on the hand feels cool because its evaporation takes away heat.

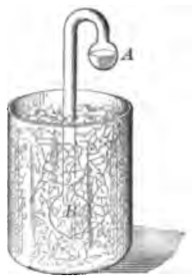


Fig. 260

Wollaston's cryophorus (Fig. 260) was designed to freeze water by its own evaporation. It contains no air, but only water and water vapor. The water is collected in bulb A, while B is surrounded with a freezing mixture. The vapor condenses rapidly in B, and the lowering of the vapor pressure causes equally rapid evaporation of the water in A. Heat is thus carried by the vapor from A to the freezing mixture until the water in A freezes.

Water may easily be made to *boil until it freezes.* A thin flat dish is supported over a broad shallow glass vessel containing strong sulphuric acid (Fig. 261), and the whole is inclosed in a low receiver on an air pump. The success of the experiment depends on removing the water vapor so rapidly that the boiling point is reduced to the freezing point. A good air pump is needed and the volume of the receiver should be no greater than is needed to cover the two dishes. When the air is exhausted, the pressure is reduced until the boiling point falls to the temperature of the water. The water then boils violently; if the vapor is removed rapidly enough by the pump and by absorption by the acid, the pressure may be reduced below 4.6 mm. of mercury, and the boiling will then continue until the water freezes. The heat to do the internal work of evaporation is drawn from the thin dish and the water itself.

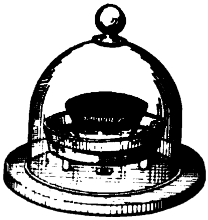


Fig. 261

Very low temperatures may be produced by the vaporization of more volatile substances. Thus, by opening a small orifice in a strong cylinder containing liquid carbon dioxide under great pressure, the spray escaping into a wooden box lined with asbestos cools so rapidly by evaporation that some of it freezes in the form of fine white snow at a temperature of – 79°.

The property of the absorption of heat by liquids evaporating at a low temperature is applied in ice machines. Ammonia, for example, is first condensed by pressure and cooling to a liquid with about one-tenth of its weight of water. It is then evaporated under reduced pressure obtained by vacuum pumps, and its temperature falls low enough to freeze water in vessels about it or submerged in it. The process is made continuous by returning the gaseous ammonia to a condensing chamber cooled with water. It thus passes repeatedly through the same cycle of changes.

410. Aqueous Vapor in the Atmosphere. — The atmosphere always contains aqueous vapor, and if it is sufficiently cooled

a temperature will be reached at which the pressure of the aqueous vapor equals the saturation pressure. The vapor then condenses in the air as rain or snow and on the surface of bodies as dew or hoar frost. This temperature is known as the *dew point.*

Condensation commences about nuclei, and in the absence of these the air may be supersaturated with moisture. Each minute mote of dust floating in the air serves as such a nucleus. The more numerous the nuclei, the smaller the drops, and small drops sink through the air very slowly on account of surface friction. A fog is composed of much smaller drops than those falling as rain. The prevalence of fogs in large cities in a moist climate is accounted for by the presence in the air of innumerable particles of dust and soot. A rain clears the air by dragging down these particles weighted with moisture.

It is now known that air ionized by ultra-violet light, by Roentgen rays, or by radioactive substances, also furnishes nuclei for condensation (§ 675).

Condensation of aqueous vapor may readily be shown by passing a beam of strong light through a large glass receiver on an air pump in a darkened room. With only moderately moist air, a single stroke of the pump produces a thick cloud of condensed vapor, showing splendid iridescent diffraction effects. The expansion of the air under pressure cools it below the dew point. The condensed water is in a state of fine division.

411. Relative Humidity. — Air is said to be damp when it is nearly saturated with water vapor. Since the saturation pressure rises rapidly with the temperature (Fig. 258), the heating of the air, while the quantity of aqueous vapor in it remains unchanged, removes it farther from the saturation point and diminishes its dampness. A cubic meter of air saturated at 0° contains 4.87 gm. aqueous vapor ; at 15° it contains 12.76 gm. ; and at 30°, 30.15 gm. Therefore if air which is saturated with moisture at 15° is heated to 30°, its humidity will be only about two fifths the amount required

for saturation. Thus, when damp air from outdoors passes through a hot-air furnace, it becomes dry, not because it has lost any aqueous vapor, but because its capacity to take up this vapor has been greatly increased by the rise of temperature. In winter the humidity is usually greater than in summer because the temperature is lower, and the amount of aqueous vapor necessary to saturate the air is less.

Relative humidity is the ratio of the pressure of the aqueous vapor present at a given temperature to the saturation pressure at the same temperature. It is measured by determining the actual pressure of the aqueous vapor in the air and comparing it with the maximum pressure at the same temperature obtained from tables. This is the method applied by all dew point instruments, which are called *hygrometers*.

412. Regnault's Hygrometer.—The superior form of hygrometer devised by Regnault consists of two thin polished silver thimbles, A and B (Fig. 262), into which are fitted glass tubes open at both ends. The tube A is half filled with sulphuric ether and is closed with a stopper through which pass a thermometer t' and a bent tube C extending down nearly to the bottom of the silver thimble. The other tube contains only a thermometer t. The exhaust tube DE leads to an aspirator.

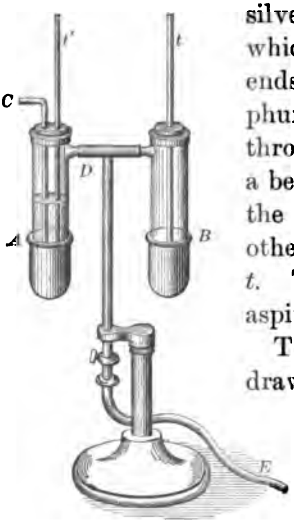


Fig 262

To make an observation, the air is drawn through C by the aspirator. It bubbles up through the ether and causes it to evaporate rapidly. The temperature of A is thus lowered; and when the dew point is reached, it is indicated by a dimming of the silver tube A as compared with B, which remains at atmospheric temperature. The thermometer t' is read as

soon as the dimming is apparent. The aspirator is stopped at the same time, and the temperature is again read at the instant when the dew disappears.

The temperature given by t' is then the dew point. The saturation pressures corresponding to both temperatures are taken from the table, and their ratio is the relative humidity. For example, if the dew point is 7° and the temperature of the air 20°, the corresponding saturation pressures are 7.49 and 17.39 mm. respectively. The actual pressure of the aqueous vapor is therefore 7.49 mm., and the pressure for saturation at 20° is 17.39 mm. Hence the relative humidity is $7.49/17.39 = 0.431$, or 43.1 per cent.

VI. LIQUEFACTION OF GASES

413. Conditions for Liquefaction. — Under atmospheric pressure a number of substances are known to us in both the liquid and the gaseous states. Water is liquid below 100° and a vapor at higher temperatures; alcohol is liquid below 78° and a vapor above; sulphuric ether is liquid below 35° and a vapor above. If we had no means of reducing the temperature below freezing, sulphur dioxide at atmospheric pressure would be known to us as a gas only, since it boils at $-8°$.

When the temperature of a substance in the gaseous state is lowered artificially, and its boiling point is at the same time raised by pressure, the two temperatures approach each other; and if these processes are carried far enough to bring the two temperatures together, liquefaction ensues.

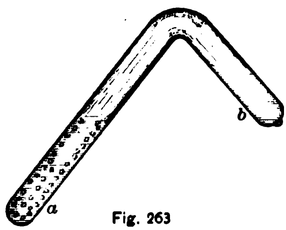


Fig. 263

Faraday was one of the first to liquefy chlorine, carbon dioxide, and ammonia, by combined cooling and pressure. His apparatus was of the simplest character, consisting merely of a stout bent tube (Fig. 263). The materials to liberate the

gas by heat were placed in this tube, and it was then sealed off. The limb *b* was surrounded by a freezing mixture, the limb *a* was heated, and the pressure was produced by the gas itself. When *a* was filled with sodium carbonate and was heated, the released carbon dioxide condensed to a liquid in *b*.

By means of a small inclosed pressure manometer it was found that the pressure in every case increased to the point where condensation set in; after that it remained constant so long as the temperature of the condensed liquid was kept the same. This pressure was that of the saturated vapor at the given temperature.

414. The Critical Temperature. — The renowned experiments of Dr. Andrews, published in 1869, established the existence for gases of a critical temperature, above which no amount of pressure produces liquefaction. The phenomena discovered by Andrews may be shown by means of a stout glass tube half full of liquid carbon dioxide and sealed before the blowpipe. Careful heating of the tube causes the liquid to expand rapidly; at the same time the pressure increases, until the surface of separation between the liquid and the vapor becomes less clearly marked, and finally the two phases merge into each other at a temperature of 31°. The entire tube is then filled with a homogeneous fluid. When this fluid has cooled a little, a thick cloud suddenly makes its appearance on the top, and the surface of separation between the liquid and the vapor is again visible. The temperature at which the liquid surface disappears or reappears is the *critical temperature.* Above its critical temperature a gas cannot be liquefied by any pressure, however great. Above the critical temperature of 31° the distinctions between liquid and gaseous carbon dioxide cease to exist.

If the pressure on the gas above 31° be increased to 150 atmospheres, the volume will diminish, but there will be no sudden decrease at any point. If now the temperature be

gradually lowered to about 20°, the fluid is then clearly a liquid. The substance has passed from the gaseous to the liquid form by a continuous process and without any sudden evolution of heat. A gas and a liquid are then only widely separated forms of the same condition of matter, and the passage from one to the other may be made without breach of continuity.

The phenomena attending the condensation of carbon dioxide may be shown by plotting pressures and volumes as coördinates. Each curve (Fig. 264) corresponds to a single temperature and is therefore called an *isothermal line.* The coördinates of the point P denote the volume and pressure at a temperature of 21.5°. The volume decreases to a point A with increasing pressure. At A condensation sets in and continues along the line AB with constant pressure until the whole mass is liquefied. A further increase of pressure produces but a slight change in volume along the line BC.

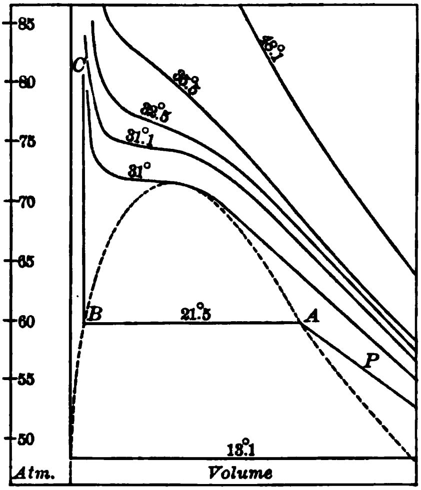


Fig. 264

The region below the dotted line, which passes through all such points as A and B, is one in which the liquid and the vapor exist together in equilibrium. Above this line the substance is either wholly a liquid or a gas. The top of the dotted curve corresponds to the critical temperature, and above it the isothermals show only a point of inflection. The isotherm for 31° touches the dotted line where the former is flattened.

The following are the critical temperatures for several substances: —

Hydrogen	– 220°	Chlorine	130°
Nitrogen	– 146	Sulphur dioxide	155
Air	– 141	Ether	197
Oxygen	– 118	Acetone	246
Fluorine	– 120	Alcohol	239
Carbon dioxide	31	Carbon bisulphide	273
Ammonia	130	Water	365

415. Distinction between a Gas and a Vapor. — The discoveries of Andrews permit of drawing a distinction between a gas and a vapor. By a *vapor* is meant a substance in the aëriform state at a temperature below the critical point. A *gas* is a substance in the aëriform state above its critical temperature. A vapor can be converted into a liquid by pressure alone, and can exist in contact with its own liquid. A gas cannot be liquefied by pressure alone, but only by first reducing its temperature below the critical temperature and thus converting it into a vapor. Below 31° carbon dioxide may exist as a vapor; above 31° it is a gas and cannot be liquefied. At the critical temperature the heat of vaporization becomes zero.

416. Gases with Low Critical Temperatures. — Oxygen, nitrogen, and hydrogen resisted liquefaction for several years after the discovery of a critical temperature. The critical temperature of these gases is below the temperature of solid carbon dioxide.

In 1877 Cailletet and Pictet liquefied oxygen and nitrogen by resorting to a two-cycle process of cooling by evaporation under reduced pressure, and then by the cooling effect of the sudden expansion of a gas under great pressure. Oxygen under a pressure of 500 atmospheres was surrounded by liquid carbon dioxide, and this in turn by liquid sulphur dioxide. Both of these cooling liquids were rapidly evaporated under low pressure by exhaust pumps. The sulphur

dioxide cooled by evaporation withdrew heat from the liquid carbon dioxide, and the rapid evaporation of the latter finally reduced its temperature and that of the oxygen to $-130°$. At this stage the pressure of the oxygen fell to 320 atmospheres, indicating some liquefaction. It was then allowed to issue from a stopcock with great violence, and its sudden expansion reduced its temperature to such an extent that some of it was collected in the liquid state.

Later Sir James Dewar employed two circuits of liquid nitrous oxide and ethylene in successive cycles for cooling oxygen and nitrogen under great pressure. The sudden expansion of oxygen as it rushed out of a small orifice cooled it finally to the liquefying point. It was mixed with some solid carbon dioxide from the air, from which it was freed by filtration through an ordinary filter paper. It has a delicate sky-blue color, and its temperature when evaporating under atmospheric pressure is $-182°$. Nitrogen was liquefied by Dewar in the same manner, though its critical temperature is lower than that of oxygen.

417. The Regenerative Process.—The process of Linde and of Hampson, introduced in 1895, differs from the older method in leading the gas cooled by sudden expansion back through spiral tubes around the pipe containing the gas flowing out under great pressure. This method is known as the regenerative process. The gas to be liquefied is put under a pressure of 150 to 200 atmospheres by a compressor, cooled by some means to a low temperature, and then allowed to escape into a chamber in which the pressure is kept down to a few atmospheres. The sudden expansion of the gas cools it, and the cooled gas passes back through spiral tubes around the compressed gas flowing toward the expansion orifice. The gas is thus further cooled before the expansion, and this reduction continues until the temperature falls below the critical point and liquefaction begins in the low-pressure chamber. After this stage is reached, the process is continuous.

In 1898 Dewar liquefied hydrogen by cooling it under a pressure of 180 atmospheres to − 205° by means of a bath of liquid oxygen boiling under reduced pressure, and then allowing it to escape into a triple-walled silvered vacuum flask at atmospheric pressure. The internal silvering and the vacuum are for the purpose of heat insulation. The boiling point of hydrogen is 21° absolute scale. By rapid evaporation under a pressure of 25 mm., the liquid was reduced to a frothy solid, the density of which was 0.086. The density of liquid hydrogen is 0.07. Glass tubes containing air, when immersed in boiling hydrogen, have the pressure in them reduced to one millionth of an atmosphere by the liquefaction of the air in them.

Problems

1. Express the following temperatures in Fahrenheit degrees: the boiling point of nitrogen, − 195.5°; melting point of hydrogen, − 257°; alcohol flame, 1705°.

2. At what temperature will the reading on the Fahrenheit scale be the same as that on the Centigrade scale?

3. At what temperature will the reading on the Fahrenheit scale be double that on the Centigrade?

4. At what temperature will the reading on the Centigrade scale be double that on the Fahrenheit?

5. If a thermometer scale were marked 10° at the freezing point and 60° at the boiling point, what would 35° on this scale mean in Centigrade degrees?

6. The testing of a Centigrade thermometer shows that the freezing point is 0.6° and the boiling point 101°. What is the meaning of 50° on this scale if the tube is of uniform bore?

7. A glass flask holds 200 cm.^{3} of water at 0°. What is its internal capacity at 100°? The coefficient of linear expansion for glass is 0.0000083.

8. The density of a piece of silver at 0° is 10.5 gm. per cubic centimeter. What is its density at 100° if the coefficient of cubical expansion is 0.0000583?

9. A brass pendulum keeps correct time at 15°, but at 35° it loses 16 seconds a day. Find the linear coefficient of expansion of brass.

10. A solid displaces 500 cm.^{3} when immersed in water at 0°; but in water at 30° it displaces 503 cm.^{3}. Find its coefficient of cubical expansion.

11. If 3 kgm. of iron (specific heat, 0.11) at 95° are put into 3 l. of water at 10°, what will be the rise in temperature of the water?

12. A mass of 500 gm. of copper at 98° put into 500 gm. of water at 0°, contained in a copper vessel weighing 150 gm., raises the temperature of the water to 8.3°. Find the specific heat of copper.

13. Into a mass of water at 0° are introduced 100 gm. of ice at − 12°; 7.5 gm. of ice are frozen and the temperature of all the ice is raised to 0°. If the heat of fusion is 80, find the specific heat of ice.

14. If 1 kgm. of copper at 100° (specific heat, 0.095) be placed in a cavity in a block of ice at 0°, and if 119 gm. of ice are melted, find the heat of fusion of ice.

15. What will be the volume of air measuring 500 cm.^{3} at 0° if the temperature be raised to 273° and the pressure be doubled?

16. A liter glass flask of air at 0° is heated to 30°. How many cubic centimeters of air escape at 30°, neglecting the expansion of the glass?

17. How much gas must be collected at a temperature of 20° and 74 cm. barometric pressure to give 100 cm.^{3} at 0° and 76 cm. pressure?

18. A vessel filled with air at 0° is heated to 67°, when it is found that 20 cm.^{3} of air at the latter temperature have escaped. What was the capacity of the vessel at 0°?

19. A mass of hydrogen at 20° occupies a volume of 500 cm.^{3}. Find its volume at 100°, the pressure remaining the same.

20. The area of Lake Erie is 9900 sq. mi. If the density of ice is 62.3 lb. per cubic foot, and the heat of combustion of anthracite coal is 14,000 B. T. U., how many tons of coal would be required to furnish by combustion enough heat to melt the ice ½ ft. thick over the entire surface of the lake?

CHAPTER XIV

TRANSMISSION AND RADIATION OF HEAT

I. CONDUCTION

418. Three Modes of Distribution of Heat. — The distribution of heat between the different parts of a body or of a system, differing in temperature, takes place by three distinct modes. These are :

1. *Conduction*, in which heat is transmitted from one portion of matter to another in contact with it and at a lower temperature. This is a slow process, depending on differences of temperature and on the nature of the conducting substance.

2. *Convection*, in which heat is carried from one place to another by sensible masses of matter. In this manner buildings are heated by the circulation of hot water, and heat is conveyed by hot air and by steam.

3. *Radiation*, in which the energy is transferred from one body to another by the same physical process as the one by which light is transmitted, and without heating the intervening medium. By the first two modes heat is distributed through the agency of matter; in the third the ether is the medium of transmitting it as radiant energy.

419. Conduction in Solids. — When the molecules of a solid are agitated by the motion of heat, their oscillations impart similar motions to adjacent molecules. The slow transmission of this motion of heat from particle to particle is conduction.

Conduction tends to establish equilibrium of temperature. If one end of an iron rod be placed in the fire of a forge, the

other will in time become hot. In this mode of distribution heat is handed on from the hotter to the colder parts of a body by a slow process.

Different substances possess the capacity of conducting heat in very different degrees. Metals are the best conductors, while glass, wood, fire clay, wool, and feathers are poor conductors. If a sheet of writing paper is tightly wrapped around a cylinder of uniform diameter, half brass and half wood (Fig. 265), a Bunsen flame may be applied for a short time without charring the paper in contact with the brass, while around the wood it will be scorched. The metal conducts away the heat so rapidly that it keeps the temperature of the paper in contact with it below the point of ignition.

Fig. 265

A Norwegian cookstove (or fireless cooker) is a box heavily lined with felt or other poor conducting material. In it fits a metallic dish with a cover. The food to be cooked in water is first boiled in the dish in the usual way and is then transferred to the box and is inclosed in it. The conductivity of the felt and the imprisoned air is so poor that the heat is retained, and in three hours the temperature falls not more than 10° or 15° and the cooking is completed without further heating.

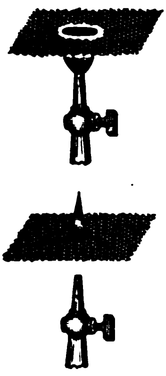


Fig. 266

The good conductivity of metal gauze was utilized by Davy in his safety lamp for miners. The flame is completely inclosed in metal and fine wire gauze. By conducting away heat the gauze keeps any fire damp outside the lamp below the point of ignition and prevents explosions. The action of the gauze is easily illustrated by holding it over the flame of a Bunsen burner (Fig. 266). The flame does not pass through unless the gauze is heated to redness. If the gas is first allowed to stream through the gauze, it may be lighted on top without being ignited below.

420. Coefficient of Thermal Conductivity. — Let AB and CD (Fig. 267) be two parallel surfaces of a homogeneous body, the thickness of which is l, one of the surfaces being maintained at a temperature t and the other t'. The length AB may represent the temperature t, and the length CD, the temperature t'. Then the temperature gradient $(t - t')/l$ is represented by the slope of the line BD, that is, by tan BDG. The quantity of heat transmitted in time T through any cross-section S of the plane EF and at right angles to the two surfaces is proportional to the area S, to the temperature gradient, and to the time, or

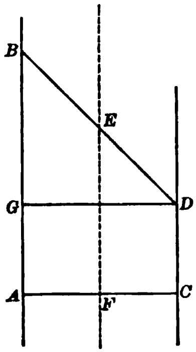


Fig. 267

$$H = KS\frac{t - t'}{l}T. \qquad (78)$$

The proportionality factor K, which depends on the nature of the substance, is the *coefficient of thermal conductivity*. It is the *time rate with which heat is transmitted through unit area when the temperature gradient is unity.*

If the temperature is measured in Centigrade degrees, the dimensions in centimeters, and the time in seconds, the quantity of heat is in calories.

Practical methods of measuring thermal conductivity are not applied to such a body with parallel sides, but to the flow of heat along a bar, one end of which is maintained at a constant temperature, while the other is at the temperature of the room. The temperature gradients are then represented by the tangents to a curve obtained by measuring the temperatures at equal distances along the bar. The heat flowing past any cross section of the bar is all dissipated from the surface beyond the section. The relative conductivities of two bars can be determined by getting their temperature gradients; but to measure the coefficient K, another experiment is necessary in order to find the rate of cooling. Then the total quantity of heat traversing any section of the bar may be calculated.

421. Conductivity in Wood and Crystals. — The investigation of many kinds of wood shows that heat is conducted

better along the fibers than across them. Further, the conductivity perpendicular to the fibers and to the ligneous rings is greater than in a direction tangential to them. The conductivity in the first of these three rectangular directions is from two to four times as great as in the last.

A similar difference of conductivity has been found in the case of laminated rocks, the conductivity being better along the planes of cleavage than across them. The same statement may be made with respect to bismuth.

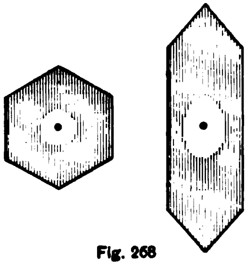


Fig. 268

If two plates be cut from quartz crystals, one perpendicular to the optic axis and the other parallel to it, and if a minute hole be made through each plate for the passage of a fine wire which may be heated by an electric current, then a film of wax on the crystal will be melted in the form of a circle when the section is at right angles to the axis, and in the form of an ellipse when the section is parallel to the axis (Fig. 268). Quartz and calc-spar conduct heat best along the axis and equally well in all directions perpendicular to it; tourmaline conducts best at right angles to its axis.

422. Conductivity of Liquids and Gases. — Liquids have low coefficients of conductivity as compared with solids. Gases are still poorer conductors of heat. If liquids or gases are heated at the bottom, the heat distributed by convection masks any distribution by conduction. This difficulty is overcome in part by heating at the top, but the results are still complicated with diffusion and with conduction by the containing vessel.

All liquids except molten metals are poor conductors. The upper strata of water in a test tube may be boiled for some time without melting a lump of ice confined at the bottom of the tube. Support a simple air thermometer by

a ring stand (Fig. 269), and cover the bulb to a depth of a centimeter with water; pour a spoonful of ether on the water and set it on fire. The index of the thermometer will show that little if any heat is conducted to the bulb. So feeble is the flow of heat through liquids that the results of measurement are always open to the suspicion that the transport has been brought about by diffusion and convection.

Fig. 269

The difficulties encountered in measuring the conductivity of liquids are exaggerated in the case of gases, so that they become almost insuperable. Many facts, however, go to show that heat is conveyed very imperfectly by gases, except under conditions favorable to convection. The interstices filled with air in bodies made of wool, hair, feathers, or fur, make them poorer conductors than they are after the air spaces have been diminished by compression. Some solids which conduct fairly well are very poor conductors when reduced to a powder. The solids are made discontinuous by the introduction of air.

II. CONVECTION

423. Convection in Liquids. — The transmission of heat by convection is accomplished by the translatory motion of heated matter. Convection of heat by currents of warm water may be illustrated by heating a large beaker of water containing some bits of cochineal. A stream of warm water ascends along the axis above the burner, and currents of cooler water descend along the sides. In Faraday's apparatus to illustrate convection currents (Fig. 270), the flask and connecting tubes are completely filled with water up to

a point above the open end of the vertical tube *C*. The water begins to circulate as soon as heat is applied to the flask by means of a Bunsen burner. To make the circulation visible, the liquid in the flask may be colored red with some aniline dye, and that in the reservoir at the top blue.

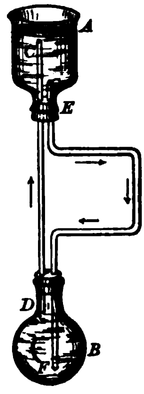


Fig. 270

This experiment illustrates the method of heating buildings by hot water. A pipe rises from the top of the boiler to an expansion tank in the upper part of the building. From this tank the hot water is distributed through the several radiators, and finally enters the boiler again at the bottom. The water loses heat in the radiators of large surface, and becomes denser. The heat of the boiler and the loss by radiation and convection at the radiators produce unequal hydrostatic pressures, which give rise to continuous currents as long as the heat is applied.

The Gulf Stream is a convection current on a gigantic scale, and it transports enormous quantities of heat from the equatorial regions, and distributes it over the British Islands and the western part of the continent of Europe, thus contributing largely to their mild climate. A counter current of cold water flows south from Greenland and washes the Atlantic coast of America. Hence the contrast between the climate along the Hudson and the Tiber in about the same latitude.

424. Convection in Gases. — Convection currents are more easily set up in gases than in liquids. The heated air over a flame rises rapidly, and its place is taken by a lateral inrush of cold air. The current of hot air over a gas flame or a bar of hot iron may easily be shown by placing them in the path of a beam of light from a projection lantern. The wavering outlines of the rising stream of air can be distinctly seen because the refraction for the hot air is less than for the cold. When the atmosphere is in unstable thermal equilibrium, especially over heated sand, a telescope focused on a vertical straight object gives a wiggling line as an image.

Convection air currents on a large scale are present near the seacoast. The wind is a sea breeze during the day,

because the air moves from the cooler ocean to take the place of the air rising over the heated land. As soon as the sun sets, the ground cools rapidly by uncompensated radiation, and the air over it is cooler than over the sea. Hence the reversal in the direction of the wind, which is now a land breeze.

The trade winds are similar currents on a still larger scale. The earth and the air are highly heated by the vertical rays of a tropical sun ; the air expands, rises, and overflows northward and southward. The denser air from both hemispheres flows in to take the place of the ascending mass. The air currents toward the equatorial belt lag behind the west-east rotating earth, and blow as northeast and southeast trade winds.


Fig. 271

425. Convection by Hydrogen. — The rapid convection of heat by hydrogen gas formed the subject of a celebrated experiment by Dr. Andrews. A thin platinum wire, which could be heated by an electric current, was stretched along the axis of a glass tube. Inlet and exhaust tubes were provided for filling and exhausting (Fig. 271). When the tube was exhausted of air, the current was adjusted so as to heat the wire to vivid brightness without fusing it. The introduction of air sensibly diminished the brightness of the wire ; but when the tube was filled with hydrogen, the wire was scarcely red hot. In an atmosphere of hydrogen the light and rapidly moving molecules carry frequent charges of heat from the wire to the cooler walls of the tube.

The incandescent lamp is made with the filament in a high vacuum to avoid the loss of heat by convection. For this reason even an inert gas, like nitrogen, cannot be used to fill

the bulb, because the energy of the heated filament is then rapidly conveyed from it to the glass bulb, and heats it at the expense of the brightness of the filament. The bulb heats to a still higher temperature when the filament is inclosed in an atmosphere of hydrogen. Very hot bulbs indicate lamps with too low a vacuum.

426. Ventilation. — The office of a chimney is to produce a convection current of air. The heated air in a chimney rises because it is lighter than the air without. The external pressure is therefore only partly counterbalanced by that of the air in the flue. If the chimney happens to be colder than the external air, there is a down draft, or the chimney smokes.

The office of a lamp chimney is to increase the supply of oxygen to the flame. The air within it is heated by the flame and rises; at the same time cold air flows in at the bottom to restore the equilibrium.

Place a lighted candle at the bottom of a tall lamp chimney. Ingress of air may be prevented by pouring a little water into the outer dish (Fig. 272). The flame soon goes out for lack of oxygen. If a T-shaped partition be now placed in the chimney and the candle relighted, it will continue to burn. If a piece of smoldering brown paper be held over the tube, the smoke will descend on one side of the partition and ascend on the other. This is a true convection current, supplying oxygen to the candle and carrying off the products of combustion.

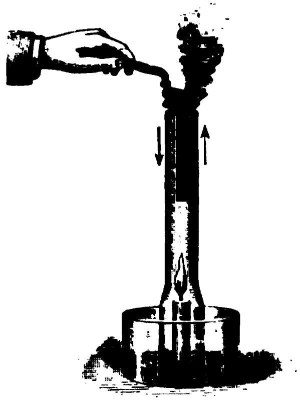


Fig. 272

III. RADIATION AND ABSORPTION

427. Cooling by Radiation. — A hot body imparts its energy of heat to other bodies, not only by conduction and by convection currents, but also by *radiation* without the agency of any intervening material. We feel the heat radiated from the sun or from a hot stove and warm ourselves before an open grate fire. In the best vacuum of an incandescent lamp bulb, the hot filament cools rapidly by radiation unless the energy is supplied as fast as it is dissipated.

The transmission of the energy given out by a hot body through a medium without heating it is a physical operation identical with the transmission of light. The experimental evidence of the physical identity of "radiant heat" and light is overwhelming. It is reflected, refracted, polarized, and is subject to interference the same as light. This energy during transmission is not heat, but it is transformed again into heat by absorption in bodies which do not transmit it. Light waves, heat waves, and electric waves are essentially identical except in the length of the waves, and they all travel with the stupendous velocity of 300 million meters a second. The light effects, the heat effects, and the chemical effects of radiation are only different aspects of the same physical phenomena.

428. Law of Cooling. — Newton's law of cooling is that the rate is proportional to the difference of temperature between that of the hot body and the inclosing chamber. This law represents the facts fairly well when the difference of temperature does not exceed a few degrees, but it fails notably when the difference is as much as 50°.

The law of Stefan, discovered in 1879, is that the *rate of emission* is proportional to the fourth power of the absolute temperature. The heat actually lost by a body is the difference between what it radiates and what it receives from surrounding bodies. Then by Stefan's law the loss of heat R

per square centimeter per second may be expressed as

$$R = c(T_1^4 - T_2^4), \tag{79}$$

where T_1 is the absolute temperature of the hot body and T_2 that of the inclosure. The constant c for a perfectly black body has been found to be about 1.28×10^{-12} calories.

When $T_1 - T_2$ is only a few degrees, Stefan's law is equivalent to Newton's law of cooling.

429. The Radiometer. — Among the instruments which have been used to measure the intensity of radiation is a modification of Crookes's *radiometer* due to Professor E. F. Nichols. Sir William Crookes invented the radiometer in 1873 while investigating the properties of highly attenuated gases. It consists of a glass bulb exhausted to a pressure not exceeding 7 mm. of mercury (Fig. 273). Within the bulb is a light cross of aluminum wire carrying small vanes of mica, one face of each being coated with lampblack; the whole is mounted so as to revolve lightly on a vertical pivot. When the radiometer is placed in sunlight, or receives the radiation from a hot body, the cross revolves with the blackened faces of the vanes retreating from the source of radiation.

Fig. 273

The explanation of this interesting instrument is found in the kinetic theory of gases that the mean free path of the molecules between collisions with other molecules, at this low pressure, becomes equal to the distance between the vanes and the wall of the bulb. The infrequent collisions among the molecules in such a vacuum prevents the equalization of pressure throughout the tube. The blackened side of the vanes

absorbs more heat than the other, and the gas molecules rebound from the warmer surfaces with greater energy than from the opposite ones, and thus give the vanes an impulse as a reaction the other way. This impulse is the equivalent of a pressure, but the residual gas has lost the power of rapid adjustment of pressure throughout the bulb. When the vacuum is not good, the increased energy of the molecules projected from the blackened vanes is so rapidly distributed that the differential pressure on the two sides of a vane becomes evanescent.

In the Nichols instrument the rotating system is delicately suspended by a quartz fiber, and the radiation is admitted through a window of fluorite, which is remarkably transparent to radiations of all wave lengths. The suspended system carries a very light mirror, and the deflections are read by means of a telescope and scale.

430. Reflection of Invisible Radiation. — Aside from the simple observation (as shown by a fire screen) that radiant heat travels in straight lines like light, the most obvious analogy between the two is found in their common obedience to the law of reflection.

Two large concave mirrors are placed several meters apart, as in Figure 274. The two may be adjusted in position by

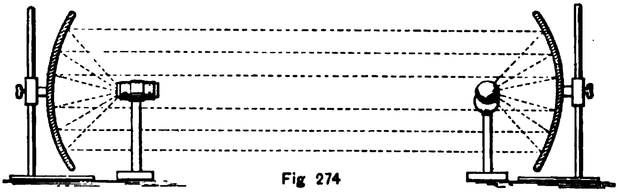


Fig 274

means of a candle placed at the focus of one of them. The position of its image may be marked at the focus of the other mirror. Then if the candle be replaced by a heated iron ball, and if the blackened bulb of a thermometer, or the blackened face of a thermopile (§ 573), be placed at the marked focus, either instrument will show that the radiant heat is reflected to the same focus as the light of the candle. The

reflection of the non-luminous radiation from the two mirrors takes place in the same manner as that of the luminous radiation, for both converge to the same point. If the ball be heated to a dull red, the convergence of the heat at the focus may be felt by the hand. The thermopile will detect it when the ball has cooled to such an extent that it may be held in the fingers.

If the ball be replaced by a piece of ice, a delicate thermopile and galvanometer will show that the thermopile is cooled. In this case the thermopile radiates more heat to the ice than the ice radiates to it. It therefore cools.

431. The Law of Inverse Squares. — Melloni was the first to perform an ingenious experiment to demonstrate that the invisible heat radiation received from any small area varies inversely as the square of its distance from the source. *BC* (Fig. 275) is a tank filled with hot water and coated on its front side with lampblack. Let the thermopile with a converging cone be placed in the position *A*, and let the deflection of the galvanometer be noted. Then let the thermopile be moved to double the distance from the tank at *A′*. The galvanometer will indicate the same current as before. The radiating surfaces in the two cases are the bases of the dotted cones. Their linear dimensions are as one to two and their areas as one to four. Since the radiation from a fourfold area produces the same effect at twice the distance, the intensity of the radiation received from any small area must vary inversely as the square of the distance. In the same way a uniformly red-hot surface, viewed through a tube, appears equally bright at all distances, so long as the surface fills the field of view through the tube.

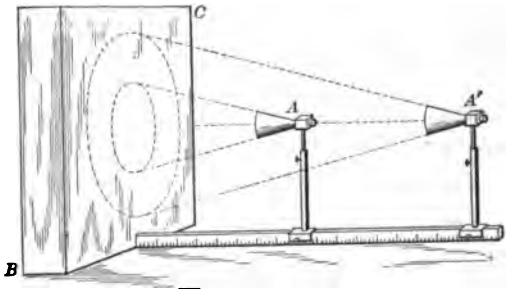


Fig. 275

432. Refraction of Heat Radiation. — Herschel observed dark heat radiation in the solar spectrum beyond the red end, or

the invisible infra-red rays. They are refracted and dispersed along with radiation of shorter wave length from a source of high temperature like the sun.

Melloni discovered that rock salt transmits all wave lengths of non-luminous radiation with nearly equal facility, while almost every other substance absorbs them with avidity. Clear glass is as opaque to radiation from a non-luminous source as black glass is to visual radiation. By employing rock salt prisms and lenses Melloni demonstrated that the radiation from a body at a temperature as low as 100° may be refracted and converged to a focus the same as light.

It has also been shown that the refrangibility decreases with the temperature of the source, and that obscure heat rays are of lower refrangibility, or longer wave length, than visual rays.

433. Absorption of Radiation.—When luminous radiations are incident on a body, in general, one portion is reflected, another is transmitted, and the remainder is absorbed. Thus, a piece of red glass reflects a part of the incident beam, transmits only wave lengths near the red end of the visible spectrum, and absorbs the rest, converting its energy into heat. If the transmitted portion is reduced to zero, the body is opaque; if the surface is covered with lampblack, the reflected light is sensibly zero and the entire incident beam is absorbed. The absorption that rejects the red only is called selective absorption, while that of lampblack is general.

This division of the incident radiation, either by general or selective absorption, is not peculiar to those radiations that affect the eye. Bodies which transmit radiant heat are said to be *diathermanous*, while those that absorb it are called *athermanous*. A body transparent to light is not on this account transparent to non-luminous radiation. Clear glass is transparent to radiation somewhat beyond the violet of the spectrum, but it is very athermanous to long heat waves.

If a sheet of glass be held between the heated ball and the mirror in the experiment of Figure 274, little or no heat will be detected at the focus of the distant mirror. All glass exhibits selective absorption, and colored glass has its range of absorption extended to some portions of the visible spectrum. In a physical sense clear glass itself is not colorless.

Hard rubber in thin sheets is opaque to light, but transparent to long heat waves. Carbon disulphide transmits in almost equal degree the luminous and the non-luminous rays; but if iodine be dissolved in it, more and more light will be cut off as iodine is added, until at length the solution becomes opaque. But the solution is still diathermanous. Tyndall inclosed it in a hollow lens with rock salt faces and showed that it transmits enough heat from an electric arc light to raise platinum to incandescence at the focus.

Such facts as these lead to the conclusion that selective absorption is not peculiar to the visible spectrum. There are Fraunhofer lines, or gaps in the continuity due to absorption, in both the infra-red and the ultra-violet, as well as in the visible spectrum. In the case of liquids the absorption bands in the spectrum are fairly broad; with gases and vapors as the absorbing media, the bands are sharply defined and consist of a number of narrow lines in different parts of the spectrum.

434. Theory of Exchanges.—If a warm body, such as a thermometer, be hung in an inclosure cooler than itself, it will lose heat, and even in a vacuum thermal equilibrium will at length be reached by radiation alone. Does all radiation cease when the body and the inclosure are at the same temperature, and does the body radiate no heat when surrounded by an inclosure warmer than itself? A cold body introduced into the inclosure would at once receive heat by radiation; but its presence can have no effect on the radiation of other bodies within the same inclosure.

Prevost came to the conclusion many years ago that radia-

tion continues all the time, and that its rate has no relation to the temperature of other bodies in the inclosure, but is a function of the nature of its surface and of its temperature.

In the experiment of Figure 274 with the mirrors, it would be unscientific and unnecessary to suppose that cold is radiated by the ice. The thermopile radiates toward the ice exactly as it does toward the hot ball, but it receives from the ice less heat than it loses by radiation, and its temperature therefore falls.

The two processes of radiation and absorption go on together, and there is a continuous interchange of energy between bodies by radiation. Heating and cooling depend on the differential effect of the radiation emitted and absorbed. A stationary temperature is maintained only so long as the emission and the absorption balance each other.

435. Absorptive and Emissive Power. — A body that absorbs all radiation incident on it is known as a "perfectly black" body. The incident energy is all transformed by such a body into heat. Lampblack is an approximately "black" body.

The *emissive power*, or *emissivity*, of any surface is defined as the ratio of the energy emitted by it to that emitted by a black body of equal area at the same temperature and in the same time.

The *absorptive power* of a surface is the ratio of the energy absorbed by it to that absorbed by a black body of equal area in the same time.

Since a black body absorbs all radiation incident on it, the absorptive power of a body may be more simply defined as the fraction of the whole incident energy which it absorbs.

Emissive and absorptive powers are connected by the simple relation, that for any given surface at a given temperature, *the two are equal to each other*.

436. Equilibrium of Radiation. — The equality of the emissive and the absorptive powers, which flows from the principle of Prevost, applies even to such specific differences as wave

length and polarization of radiation, whether luminous or non-luminous.

Sodium vapor is a remarkable illustration. When heated it emits radiations of two wave lengths differing but little from each other. Now, if light from a white-hot solid is transmitted through relatively cool sodium vapor, the spectroscope reveals two dark absorption lines identical in position with the two bright lines which self-luminous sodium vapor emits. The energy absorbed by the cooler sodium vapor is greater than the energy emitted; but if its temperature be raised to that of the source of the radiation, the emissivity will equal the absorptive power and the absorption lines will disappear.

If the absorptive and emissive powers are equal, then good reflectors are poor radiators. A pot of red-hot lead examined in the dark is more luminous where the surface is covered with dross than where it is clear.

If a piece of platinum foil, with a figure drawn on it in ink, be heated in a Bunsen flame held under it, the blackened portion will be more highly luminous than the rest when viewed from the tarnished side. If it be viewed from the reverse side, the figure in ink will be darker than the adjacent surface. Since the tarnished surface radiates more than the clean surface, it is cooler and appears dark by contrast on the reverse side.

In a striking experiment of Stewart a piece of white stoneware with a black pattern on it was heated to white heat. In the dark, the black surface shone much more brightly than the white, so that the white and black in the pattern were curiously exchanged.

A transparent piece of tourmaline, cut parallel to its axis, absorbs nearly all the light polarized in a plane parallel to the axis. If Prevost's principle extends to polarization, such a plate when heated red-hot should emit light polarized in the same plane as the light which it absorbs. This proved to be true.

Again, any incombustible colored body in a bright coal fire does not alter the color of the light emitted after it has attained the temperature of the hot coals. A piece of red glass, for example, transmits red from the glowing coals and radiates the greenish light which it absorbs when cold. The light which it radiates exactly compensates for the light which it absorbs.

Problems

1. What will be the resulting temperature if 5 kgm. of water at 90° are mixed with 5 kgm. of ice at 0°?

2. 4 kgm. of ice at 0° are put into 6 kgm. of water at 40°. Determine the result.

3. How much steam at 100° will be required to melt 310 gm. of ice at 0° and to raise the temperature of the water to 16°?

4. 10 gm. of steam at 100° are blown into a mixture of ice and 49.7 gm. of water at 0°. The final temperature of the water is 5°. Find the quantity of ice.

5. What is the heat of vaporization of water derived from the following data: 10 gm. of steam at 100° condensed in 610 gm. of water at 15° raised its temperature to 25°?

6. When heat was supplied at a constant rate to a certain block of tin it was found that the temperature rose 2° a second. After the melting point was reached, the temperature remained constant for 130 seconds, when all the tin was melted. If the specific heat of tin is 0.054, find its heat of fusion.

7. How many degrees would the air in a room $6 \times 4 \times 3$ m. be warmed by the condensation in the radiator of 1 kgm. of steam at 100°, if one liter of air weighs 1.29 gm.?

8. If the coefficient of thermal conductivity of iron is 0.164, how many calories will be conducted through a plate of iron 1 m. square and 0.3 cm. thick if the two sides are kept at 0° and 60° for one hour?

9. A brass plate 1 cm. thick and 100 sq. cm. in area was kept in contact with steam at 100° on one side and with melting ice on the other. If 22.9 kgm. of ice were melted in 10 minutes, what is the coefficient of thermal conductivity of the brass?

10. A plate of glass 2 cm. thick and 3×4 m. in area separates two rooms which remain at 0° and 25° respectively. If the coefficient of thermal conductivity of glass is 0.015, how much heat is given off by the glass per minute?

CHAPTER XV

THERMODYNAMICS

I. MECHANICAL EQUIVALENT OF HEAT

437. First Law of Thermodynamics. — Evidence has been accumulating that heat is a form of energy. It is a familiar fact that mechanical energy is readily converted into heat; also that in purely mechanical processes the total energy remains constant. It is also true that in a system of bodies between which only thermal processes take place, the total quantity of heat remains constant. The inquiry now relates to the constancy of energy during the reciprocal conversion of heat and mechanical work.

The first law of thermodynamics relates to this reciprocal conversion. It is the principle of the Conservation of Energy applied specifically to heat. Maxwell expressed it as follows:

When work is transformed into heat, or heat into work, the quantity of work is mechanically equal to the quantity of heat.

In symbols this relation may be written

$$W = JH, \tag{80}$$

where W is the work measured in ergs, H the quantity of heat in calories, and J the constant ratio between W and H known as the "mechanical equivalent of heat." (Obviously H and W may be measured in any suitable units.)

438. Joule's Determinations. — The investigations of Joule in support of the first law of thermodynamics and to find the mechanical equivalent of heat extended from 1842 to 1878.

They demonstrated that whenever mechanical energy is converted into heat, a definite ratio always exists between the numerical expressions for the two, whatever may be the method by which the one is converted into the other. Joule's experiments left little for subsequent investigators except refinement of details and an increase of the scale on which the experiments were conducted.

In the last experiment of Joule in 1878 water was churned with paddles, and the ratio was measured between the work expended in turning the paddles and the number of heat units generated. The work was measured by the following devices:

A calorimeter *h* (Fig. 276), containing the water, was supported on a hollow cylindrical vessel *w*, which floated in water in *v*. It was thus free to turn around a vertical axis and the pressure was taken off the bearings. The paddles in the calorimeter were carried on a vertical axis *b*, about which the calorimeter could also turn. There was a horizontal flywheel at *f*, and the paddles were turned by the handwheels *d* and *e*.

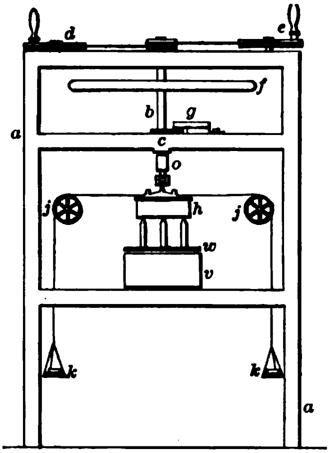


Fig. 276

To prevent the calorimeter turning by the friction between the paddles and the water, two thin silk strings were wound in a groove around it, and carried weights *k*, *k*, beyond the pulleys *j*, *j*. These weights were adjusted until they remained stationary, while the shaft and paddles revolved at a suitable uniform speed, which was recorded by the counter *g*. The weights then gave the torque necessary to keep the calorimeter at rest, that is, a torque equal to the moment of the force exerted by the paddles on the water. To measure the work done, it was

only necessary to multiply this moment by the angular velocity of the shaft. The arrangement was in fact equivalent to a friction dynamometer (§ 81).

Let m be the mass of each weight, r the radius of the groove in the calorimeter, and n the number of rotations a second. Then the moment of the two weights is $2\,mgr$, and the angular velocity is $2\,\pi n$. The former is the torque T and the latter ω. Hence in n turns the work done is

$$T\omega = 2\,mgr \times 2\,\pi n = 4\,\pi nrmg \text{ ergs.}$$

If M is the mass of the water and M' the water equivalent of the calorimeter and the paddles, and if t is the rise in temperature, the heat generated is $(M + M')t$ calories. Then the ratio of the two, or Joule's equivalent, is

$$J = \frac{4\,\pi rnmg}{(M + M')t} \text{ ergs/calorie.}$$

Joule expressed his results in terms of the temperature given by a mercury-in-glass thermometer. They were later reduced by Rowland to the hydrogen scale, with the result

$$J = 4.182 \times 10^7 \text{ ergs per calorie at } 15°,$$

$$\text{or } 1 \text{ calorie} = 4.182 \times 10^7 \text{ ergs} = 4.182 \text{ joules.}$$

439. Rowland's Determination of J.—In 1879 Rowland repeated the work of Joule in a series of very elaborate and precise experiments. Rowland's plan was the same in principle as Joule's, the chief differences being that the paddles were turned by power from a steam engine, and the revolutions were recorded on a chronograph. On the same chronograph were recorded the transits of the mercury over the divisions of the thermometer. The rate at which heat was generated in Rowland's apparatus was about 50 times as great as in Joule's. His series of determinations at different temperatures showed that the specific heat of water is a minimum at about 30°.

When reduced to the hydrogen scale, the final value for J deduced from Rowland's experiments is

$$J = 4.189 \times 10^7 \text{ ergs/calorie, at } 14.6^\circ,$$

$$\text{or } 1 \text{ calorie} = 4.189 \times 10^7 \text{ ergs} = 4.189 \text{ joules.}$$

It is apparent that the numerical value of the mechanical equivalent depends on the units employed. Rowland expressed his results also in gram meters in the latitude of Baltimore. Expressed in this way

$$1 \text{ calorie} = 427.52 \text{ gram meters.}$$

The interpretation is that if the work done in lifting 427.52 grams one meter high is all converted into heat, it will raise the temperature of one gram of water one degree centigrade at 14.6°.

In terms of the English system, the equivalent is

$$1 \text{ B. T. U.} = 778.1 \text{ foot pounds at Greenwich.}$$

This is equivalent to 1400.6 foot pounds per degree centigrade; that is, if the work done in lifting 1400.6 pounds one foot high in the latitude of Greenwich is converted into heat, it will raise the temperature of one pound of water one degree centigrade. Obviously 778.1 and 1400.6 stand to each other in the relation of 5 to 9.

440. Mechanical Equivalent from Specific Heats of a Gas. — The mechanical equivalent of heat was first calculated by Robert Mayer, a German physician, in 1842, on the assumption that the internal work done during the expansion of a gas is zero. It has since been shown by Joule and Thomson that, while this internal work is not zero, it is very small in the case of gases not easily liquefied.

Let v be the volume of unit mass of the gas at absolute temperature T. Since the volume is proportional to the temperature on the absolute scale (§ 381), v/T is the change in volume per degree, or the expansion, and pv/T, the product of the pressure and the small change of volume, is the work done by the gas during the expansion under constant pressure while the temperature changes one degree (§ 66).

The specific heat s_v is the number of calories required to raise the temperature of unit mass of a gas one degree when the volume is kept con-

stant; the specific heat s_p is the number of calories required to raise the temperature one degree when the pressure is kept constant and the gas expands. If there is no internal work done, the latter will exceed the former by the thermal equivalent of the work done by the gas in expanding under pressure.

Reduced to mechanical units, this difference may be expressed in the form of the equation,

$$J(s_p - s_v) = \frac{pv}{T},$$

or

$$J = \frac{pv}{T} \cdot \frac{1}{s_p - s_v}.$$

Under standard conditions of temperature and pressure, $T = 273$, $p = 1{,}013{,}250$, $v = 1/d = 1/0.001293$ gm. per cm.^{3}, $s_p = 0.2375$, $s_v = 0.1684$. Substituting in the last equation,

$$J = \frac{1013250}{273 \times 0.001293 \times 0.0691} = 4.16 \times 10^7 \text{ ergs/calorie.}$$

II. CARNOT'S CYCLE

441. Isothermal Processes. — An *isothermal process* is one in which the temperature remains constant. The simplest example of such a process is the isothermal expansion or compression of a gas obeying Boyle's law. If the state of the gas be represented by a pressure-volume (p, v) diagram, each curve drawn for one temperature, T_1, T_2, etc., and called an isothermal line, is a rectangular hyperbola (Fig. 277), for the equation is

$$pv = RT = \text{a constant (§ 382)},$$

and this is the criterion of a rectangular hyperbola.

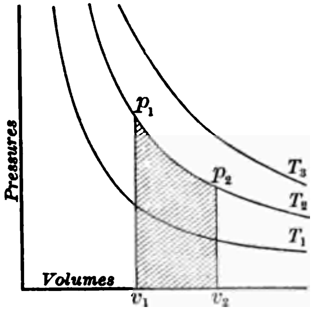


Fig. 277

If the gas of volume v_1 and pressure p_1 expands isothermally at a temperature T_1 to the state (v_2, p_2), it performs external work and absorbs heat from the outside to maintain constant temperature. The

work done during the process is measured on the scale of the diagram by the shaded area $p_1p_2v_2v_1$, included between the curve, the axis of volumes, and the two ordinates, p_1 and p_2 (§ 66).

Conversely, if the gas is compressed isothermally between the same limits, the same amount of work is done on the gas, and an equivalent quantity of heat is given out by it.

442. Adiabatic Processes. — An *adiabatic process* is one in which the working substance neither receives heat from other bodies nor loses heat to them. For a perfect gas the equation for an adiabatic process is

$$pv^{\gamma} = \text{a constant.}$$

The corresponding pressure-volume curve is an *adiabatic line.* Since work is done during an adiabatic expansion without the application of heat, the gas cools and the pressure falls below the pressure for isothermal expansion. The slope of an adiabatic line through any point is therefore steeper than that of an isothermal line through the same point.

The work done in adiabatic expansion is again measured by the area included between the adiabatic curve, the axis of volumes, and the two limiting ordinates.

The coefficient of elasticity for the isothermal expansion of a gas is numerically equal to the pressure (§ 207). But for any increment of volume in adiabatic expansion, the decrease of pressure is greater than for the same isothermal expansion. Hence the adiabatic coefficient of elasticity of a gas is greater than the isothermal coefficient.

An instance of an adiabatic process occurs in the transmission of sound through air, where the periodic compressions and rarefactions are too rapid to permit of an equalization of temperature. It is, therefore, the adiabatic coefficient of elasticity that enters into the calculation of the velocity of sound in gases (§ 208).

443. Carnot's Cycle.—To Carnot belongs the credit of introducing a *cycle* of operations, consisting of isothermal and adiabatic processes, by which external work is done and the working substance is returned to its initial state. The advantage gained by carrying the working substance, such as a gas, through a complete cycle of operations is that it neither gains nor loses internal energy. Whatever work has been done must therefore be credited to energy derived from external sources.

Carnot's cycle consists of four processes, two isothermal and two adiabatic:

1. An *isothermal expansion* from the state A to the state B (Fig. 278). The work done equals the area $ABv_1'v_1$, and heat H_1 is absorbed from without at the temperature T_1.

2. An *adiabatic expansion* from the state B to the state C. The work done equals the area $BCv_2'v_1'$. No heat is absorbed, but the temperature falls from T_1 to T_2.

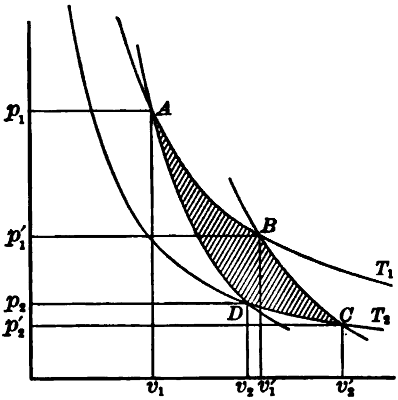


Fig. 278

An *isothermal compression* from the state C to the state D. Work is done *on* the gas equal to the area CDv_2v_2', and a quantity of heat H_2 is generated and given out at the lower temperature T_2.

4. An *adiabatic compression* from the state D to the initial state A. The work done on the gas equals the area DAv_1v_2, and the temperature rises from T_2 to T_1.

The net results of such a cycle of operations, by which the gas is returned to its initial volume, pressure, and temperature, are the following:

a. The *absorption* of a quantity of heat H_1 at the higher temperature T_1.

b. The *evolution* of a quantity of heat H_2 at the lower temperature T_2.

c. The performance of an amount of mechanical work represented by the shaded area $ABCD$ inclosed by the two isothermal and the two adiabatic lines, for

$$ABv_1'v_1 + BCv_2'v_1' - CDv_2v_2' - DAv_1v_2 = ABCD.$$

This work is obtained from the heat $H_1 - H_2$ in accordance with the first law of thermodynamics. The gas at the close of the operations is in the same state in every respect as at the beginning and has neither gained nor lost energy. It is only the agency by means of which heat is converted into work; it is therefore called the *working substance.*

444. Carnot's Engine. — Carnot's engine is an ideal one designed to show that his cycle may be used for the continued production of work from heat. Suppose the working substance D (Fig. 279) inclosed in a cylinder impervious to heat except through its bottom, which is assumed to be a perfect

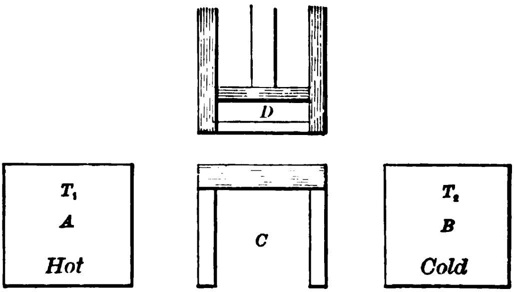


Fig. 279

conductor. A and B are two stands which are maintained at the temperature T_1 and T_2 respectively. C is a third stand, the top of which is assumed to be perfectly non-conducting. The four operations of Carnot's cycle may then be realized in the following manner:

1. The cylinder at the temperature T_1 is placed on stand A and the substance D expands isothermally. Heat flows in

through the bottom of the cylinder to maintain a constant temperature.

2. The cylinder is transferred to the stand C and the expansion is continued adiabatically until the temperature falls to T_2.

3. The cylinder is placed on the cooler stand B and the substance D is compressed isothermally. Heat flows out through the bottom of the cylinder and the temperature remains constant.

4. The cylinder is finally returned to the stand C and the substance D is further compressed adiabatically until the temperature rises to T_1.

The physical results of this cycle of operations are the same as those described in the last article. With a suitable heater A and refrigerator B, this cycle may be repeated indefinitely with continued conversion of heat into work.

These four operations may be performed *in the reverse order*. The working substance is carried around the Carnot cycle in the reverse direction, and the physical and mechanical operations are reversed. Starting with the substance in the state A (Fig. 278), the cylinder is placed on the stand C and the first expansion is adiabatic, the temperature falling from T_1 to T_2. This operation is followed by isothermal expansion with the cylinder on the stand A. Heat H_2 is absorbed at the temperature T_2. The third operation is an adiabatic compression with the cylinder on the stand C. The temperature rises to T_1 during the compression. The cylinder is finally placed on B and the substance is compressed isothermally at the temperature T_1 until it returns to its initial state at A (Fig. 278). During this compression heat H_1 is returned to the heater at the higher temperature.

The net result of this reverse order is that heat has been transferred from the refrigerator to the heater, but only at the expense of mechanical work equal to the area $ABCD$. More work has been done on the working substance than by it, and the excess has been converted into heat. Because all

the operations may be carried out in the reverse order, Carnot's engine is said to be *reversible.*

445. Efficiency of Carnot's Engine. — If temperatures are measured by a gas thermometer with a perfect gas as the thermometric substance, the quantities of heat absorbed and given out by the working substance are proportional to its absolute temperatures during the two operations of compression and expansion, or

$$\frac{H_1}{H_2} = \frac{T_1}{T_2}. \qquad \text{Then } \frac{H_1 - H_2}{H_1} = \frac{T_1 - T_2}{T_1}.$$

Efficiency of conversion is the ratio between the heat transformed into work and the heat absorbed by the working substance at the higher temperature, or

$$\textit{Efficiency} = \frac{\textit{Heat utilized}}{\textit{Heat absorbed}} = \frac{H_1 - H_2}{H_1} = \frac{T_1 - T_2}{T_1}.$$

446. Carnot's Principle. — The important principle of Carnot, derived from his reversible engine, is as follows:

"If a given reversible engine, working between the upper temperature T_1 and the lower temperature T_2, and receiving a quantity H_1 of heat at the upper temperature, produces a quantity w of mechanical work, then no other engine, whatever be its construction, can produce a greater quantity of work when supplied by the same amount of heat and working between the same temperatures."

Suppose an engine M to have a higher efficiency than a reversible engine N. Let it be coupled to N working in the reverse order. Then, since M converts a larger portion of the heat H_1 into mechanical work than N requires to restore the same quantity of heat H_1 to the source, the two engines constitute an automatic device by which M, while drawing heat H_1 from the heater, supplies to N sufficient energy to enable it to transfer from the refrigerator to the heater more heat than M withdraws. In other words, the two

coupled engines would run perpetually, transferring heat all the time from colder bodies to hotter ones. Heat might thus be collected automatically from the earth or the air and be stored in a reservoir, from which it could be drawn for the continuous performance of mechanical work. But such an operation is denied by universal experience and is inadmissible. It follows that no engine can be more efficient than the ideal reversible one of Carnot. As a practical fact no heat engine has as high an efficiency as the theoretical one of an ideal reversible engine.

Since no engine can have a higher efficiency than a reversible engine, it follows that no one reversible engine can have a higher efficiency than another. This means that the efficiency is independent of the working substance and depends alone on the temperatures between which the engine works.

447. The Second Law of Thermodynamics. — The second law of thermodynamics embodies the Carnot principle, and was stated by Clausius as follows: —

"It is impossible for a self-acting machine, unaided by any external agency, to convey heat from one body to another at a higher temperature."

Lord Kelvin expressed it in a slightly different form:

"It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects."

These statements apply only to the performance of devices working in cycles. Without this limitation, it is evident that the heat of a body, that of compressed gas, for example, may be converted into work by cooling the body below the temperature of surrounding objects; but before the operation can be repeated, the working substance must be restored to its initial condition, and this can be done only by applying energy from without.

448. Thermodynamic Scale of Temperature. — A new scale of absolute temperatures was early suggested by Lord Kelvin. Since the ratio H_1/H_2 of the quantities of heat taken in and rejected between the temperatures t_1 and t_2 depends on these temperatures alone, new numbers T_1 and T_2, having the same ratio as H_1 and H_2, might be taken to denote these temperatures in centigrade degrees, thus forming a thermodynamic scale independent of the expansion coefficient of a gas.

Such a scale leads to the conception of an absolute zero; for if the temperature T_2 is conceived to be lowered to the point where no heat is ejected by the engine, the efficiency will be unity. No lower temperature than this can exist; for if it did, by selecting this lower one as the temperature of the refrigerator, more work could be obtained from the heat absorbed than its mechanical equivalent. But such a result violates the first law of thermodynamics. Therefore there can be no negative temperature below $T_2 = 0$, which is thus the absolute zero. Kelvin's thermodynamic scale is independent of the properties of the working substance, and it has been shown to be identical with that of a perfect gas thermometer. In fact, the difference between it and the hydrogen scale between $-50°$ and $150°$ is less than $0.001°$. Even at $1000°$ the reading on the hydrogen scale is only $0.044°$ lower than on the thermodynamic scale.

449. The Steam Engine. — The most important devices for converting heat into mechanical work are the steam engine and the gas engine. In the reciprocating steam engine a piston is moved alternately in opposite directions by the pressure of steam applied first to one of its faces and then to the other. This reciprocating motion is converted into a rotatory motion by the device of a connecting rod, a crank, and a flywheel.

The working parts, in longitudinal section, of the cylinder and slide valve of a single expansion engine are shown in Figure 280. The piston B is moved in the cylinder A by the

pressure of the steam admitted through the inlet pipe *a*. The slide valve *d* works in the steam chest *cc* and admits steam alternately to the two ends of the cylinder through the steam ports at either end.

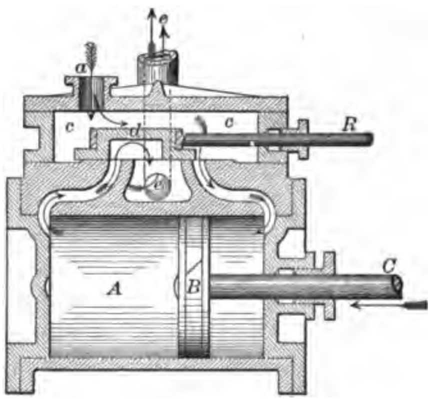


Fig. 280

When the valve is in the position shown, steam passes into the right-hand end of the cylinder and drives the piston toward the left. At the same time, the other end is in connection with the exhaust pipe *ee*, through which the expanded steam escapes, either into the air, as in a high-pressure non-condensing engine, or into a large condensing chamber, as in a low-pressure condensing engine. In the condenser the steam is condensed to water, thus reducing the back pressure on the exhaust side of the piston.

The slide valve *d* is moved through the valve rod *R* by means of an eccentric, which is a round disk mounted eccentrically on the engine shaft and has the effect of a crank. The flywheel, also mounted on the shaft, has a heavy rim with a large moment of inertia. It has, therefore, the capacity of storing enough energy during its acceleration to carry the shaft over the dead points when the piston is at either end of the cylinder. There is in the flywheel a give-and-take of energy twice every revolution, the result of which is a fairly steady rotation of the shaft.

450. The Indicator Diagram. — A steam indicator is a device for the automatic tracing of a diagram representing the relation between the volume and the pressure of the steam

in the cylinder during one stroke. This diagram is technically known as an "indicator card" (Fig. 281).

The interpretation of the successive operations, which compose the cycle and are recorded on the diagram, is the following: From *a* to *b* the inlet port is open and the pressure on the piston is the full pressure of the steam; at *b* the inlet port is closed and the steam expands in a mixed adiabatic and isothermal manner from *b* to *c*, when the exhaust port opens; at *d* the pressure is reduced to its lowest value and it remains sensibly constant during the return movement of the piston until *e* is reached, when the exhaust port closes and the remaining steam is compressed adiabatically from *e* to *f*. At *f* the inlet port opens and the pressure rises abruptly to the initial maximum, thus completing the cycle.

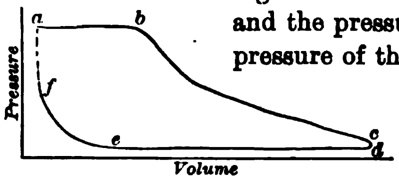


Fig. 281

The diagram bears some resemblance to the Carnot cycle, the portions *ab* and *de* corresponding to isothermal lines, and *bc* and *ea* to adiabatic lines. The work done during the stroke is represented, as in the Carnot cycle, by the inclosed area *abcdef*. It is readily calculated if the length of stroke and the scale of pressures are known.

451. The Gas Engine.—The *gas engine* is a type of *internal combustion engine*, which includes prime movers consuming illuminating gas, blast furnace gas, producer gas, gasolene, kerosene, or alcohol as fuel. The fuel is introduced into the cylinders of the engine either as a gas or as a vapor, mixed with the proper proportion of air to produce a good explosive mixture. The mixture is usually exploded at the right instant by means of an electric spark. The force of the explosion drives the piston forward in the cylinder.

In the *four-cycle* type of gas engine, the explosive mixture is drawn in and exploded every other revolution of the engine,

while in the *two-cycle* type an explosion occurs every revolution. The former type is used in nearly all motor car engines.

The operation of the four-cycle engine may be understood from a description of the four steps in the complete cycle, as illustrated in 1, 2, 3, and 4 of Figure 282. The inlet valve *a* and the exhaust valve *b* are operated by the cams *c* and *d*. Both valves are kept normally closed by springs surrounding the valve stems. The small shafts to which the two cams are fixed are driven by the small spur wheel *e* on the shaft of the engine. This wheel engages with the two larger spur wheels on the cam shafts, each having twice as many teeth as *e* and forming with it a two-to-one gear, so that *c* and *d* rotate once in every two revolutions of the crank shaft. The piston *m* is supplied with packing rings; *h* is the connecting rod, *k* the crank shaft, and *l* the spark plug.

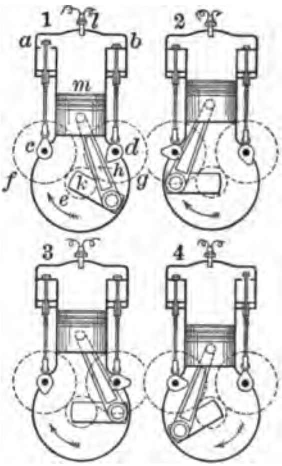


Fig. 282

The cycle is completed in four strokes, or two revolutions of the crank shaft. If the engine is running, the flywheel carries the piston down and draws in the charge through the open valve *a*, as represented in 1. In 2 both valves are closed and the piston compresses the explosive charge. Shortly after the piston reaches its highest point, the charge is ignited by a spark at the spark plug, and the working stroke then takes place, as in 3, both valves remaining closed. In 4 the exhaust valve *b* is opened by cam *d*, while *a* remains closed, and the products of combustion escape through the muffler, or directly into the open air.

A single engine is commonly constructed with two, four, or

six combined cylinders, giving one, two, or three impulses for every revolution of the shaft. The increase in the number of cylinders contributes to the steady running of the engine.

Problems

1. Victoria Falls in South Africa are 340 feet high. How much is the water heated by the fall if no heat is lost by evaporation?

2. A mass of 100 gm. moving with a velocity of 50 m. per second is suddenly stopped. If all its kinetic energy were converted into heat, how many calories would be generated?

3. An iron bullet (specific heat, 0.112) weighing 50 gm. strikes a target with a velocity of 400 m. per second. Assuming that 20 per cent of the energy of the moving bullet remains in it as heat, how many degrees will its temperature be raised?

4. How much heat is generated in stopping a train of 100,000 kgm. mass, running at 36 km. an hour?

5. How much work would be done if all the heat of combustion (7850 cal./gm.) of 1 kgm. of anthracite coal could be converted into work? It would be equivalent to how many horse power hours?

6. How much work is done against atmospheric pressure when 1 kgm. of water is converted into steam at 100°? What would be the heat of vaporization of water if this energy were not included? (At 100° the volume of the steam is 1582 times that of the water.)

7. If the heat of combustion of pure anthracite is 7844 calories per gm., find the thermal value of 1 lb. anthracite in B. T. U. Find also the maximum horse power hours obtainable from 1 lb. of anthracite if the combined efficiency of the boiler and steam engine is 25 per cent.

8. One calorie is the equivalent of 4.189×10^7 ergs. Find the number of foot pounds of energy required to raise the temperature of 1 lb. of water one degree Fahrenheit at New York, where g equals 980 cm. per second per second.

9. If a perfect reversible engine takes steam from a boiler at 153°, at what temperature must it exhaust into a condenser to have a theoretical efficiency of 40 per cent?

10. From the following data compute the indicated horse power of a steam engine, the indicator diagrams being the same on both sides of the piston: mean effective pressure, 56 lb. per square inch; diameter of piston, 10 in.; length of stroke, 12 in.; revolutions per minute, 300.

MAGNETISM AND ELECTRICITY

CHAPTER XVI

MAGNETS AND MAGNETIC FIELDS

I. PROPERTIES OF MAGNETS

452. Fundamental Facts. — Black oxide of iron, commonly called magnetite, is widely distributed and is sometimes found to possess the property of attracting iron. This property has been recognized from ancient times; it was exhibited in a marked degree by iron ores from *Magnesia* in Asia Minor, and they were therefore called *magnetic stones* and later *magnets*. They are now known as *natural magnets*, and the properties peculiar to them are called *magnetic properties*.

If a piece of natural magnet be suspended by an untwisted thread (Fig. 283), its longer dimension will point nearly in a north-and-south direction. This property of orientation, which led to the invention of the compass, has been known since the beginning of the thirteenth century; and because of this directive property the magnet in early times acquired the name of *lodestone*, or leading stone.

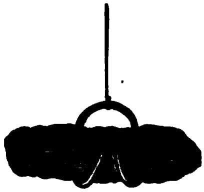


Fig. 283

Little more was known of the fundamental properties of natural magnets until Gilbert published his book entitled *De Magnete* in 1600. He named the centers of attraction near the ends of a magnet the *poles*, and the straight line joining the poles he called the *magnetic axis*.

453. Artificial Magnets. — If a slender piece of hardened steel be stroked lengthwise with a lodestone, it will acquire magnetic properties; fine iron filings will cling to it in tufts near the ends (Fig. 284); and if it be suspended, or floated on a piece of cork in water (Fig. 285), it will come to rest in a north-and-south line with the same end always pointing north. This end is called the *north-seeking pole*, and the other the *south-seeking pole.* No iron filings cling to the middle of the bar. This middle region is called the *equator.*

Fig. 284

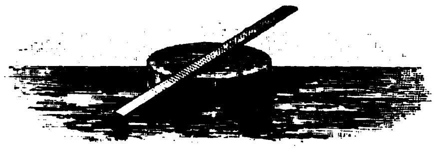


Fig. 285

A long thin rod, magnetized in the direction of its length, has its centers of attraction, or poles, very near the ends. The remainder of the magnet is apparently devoid of magnetic properties, but only apparently so. In short, thick magnets the poles are much less sharply defined.

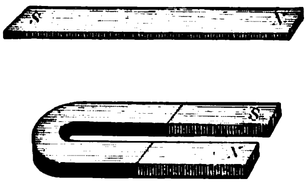


Fig. 286

Magnetized bars of hardened steel are called *artificial magnets* to distinguish them from natural magnets or lodestones. The most common forms are the *bar* and the *horseshoe* (Fig. 286), so called from their shape.

454. Magnetic Substances. — A *magnetic substance* is one capable of being affected by a magnet. Faraday showed that most substances are influenced by magnetism, but only a few show magnetic properties in a marked degree. Attraction takes place between a piece of soft iron and either pole of a magnet, but the soft iron does not retain the prop-

erty of attracting other pieces of iron, and it has no directive force when freely suspended horizontally. Neither has it fixed poles and an equatorial region.

Other substances attracted by a magnet are nickel, cobalt manganese, and chromium. Only nickel and cobalt show decided magnetic properties comparable with iron. Some gases are feebly magnetic; liquid oxygen exhibits conspicuous magnetic properties.

Another class of substances are apparently repelled by a magnet. These are called *diamagnetic*, to distinguish them from *paramagnetic* bodies like iron and nickel. Among them are bismuth, antimony, tin, copper, and some others still less strongly diamagnetic.

455. First Law of Magnetic Force. — A thin pointed bar of magnetized steel, having at the middle a cap with an inset of agate, so that the bar may turn freely on a sharp steel point around a vertical axis, is called a *magnetic needle* (Fig. 287). If the S-seeking pole of a bar magnet be presented to the N-seeking pole of a magnetic needle, they will mutually attract each other; but if the N-seeking pole of the magnet be brought near the same pole of the needle, there will be repulsion. The law of attraction and repulsion is accordingly formulated as follows:

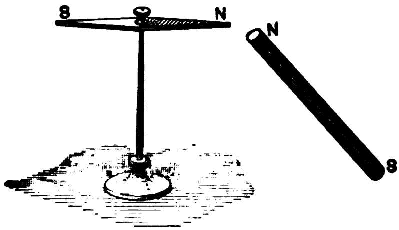


Fig. 287

Like magnetic poles repel and unlike magnetic poles attract each other.

456. Magnetic Induction. — When a magnet attracts a piece of soft iron, the iron first becomes a temporary magnet by *induction*. This piece may in turn act inductively on a

second one, and so on in a series of temporary magnets of decreasing strength (Fig. 288). But if the first piece be detached from the magnet and be slowly withdrawn, all of them will fall apart, and they will not again attract one another until they are once more brought under the inductive influence of a permanent magnet. A bar of iron near a magnet is attracted because it becomes a temporary magnet by induction, with the pole nearest to the pole of the inducing magnet of the opposite sign or name (Fig. 289). Induction thus precedes attraction.

Fig. 288

457. Permanent and Temporary Magnets. — Permanent and temporary magnets differ only in the degree with which they retain their magnetism. The softest iron retains a small amount of magnetism after it has been brought under the influence of a magnetizing force, while hardened steel retains a large proportion of it. The latter loses some of its magnetism when the magnetizing force is withdrawn, while the former loses nearly all of it. A much larger magnetizing force is required to magnetize hard steel than soft iron to the same magnetic strength. The ratio between the part lost and the part retained depends on the quality and hardness of the iron and on its after treatment. Cast iron retains an appreciable fraction of the magnetism induced in it, and this property is utilized in starting the excitation of dynamo machines. The property of resisting magnetization or demagnetization is called *retentivity*. The retentivity of hardened steel is much greater than that of soft iron.

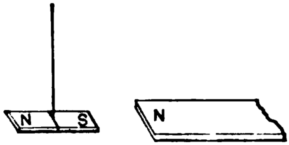


Fig. 289

458. Effect of Heat on Magnetism. — If a permanent magnet be heated to a bright red heat, all signs of magnetism

disappear. Up to 680° iron shows but a slight change in its magnetic properties; above this temperature a rapid decrease in magnetic susceptibility takes place, and at about 750° it ceases entirely to be magnetic and is quite indifferent toward a magnet. Nickel loses its magnetic properties at about 350°. Chromium ceases to be magnetic at about 500°. Manganese is magnetic at temperatures near 0° only. According to Dewar, when iron is chilled to − 200° in liquid oxygen, its susceptibility is twice as great as at 0°.

The loss of magnetization by heat in the case of nickel is beautifully shown by the simple apparatus of Figure 290, designed by Bidwell. A thin tongue of nickel is soldered to a copper disk and the whole is blackened and suspended by silk threads. A permanent magnet is held in such a position that it retains the nickel tongue just over the flame of an alcohol lamp. When the nickel reaches the temperature of about 350°, the magnet releases it, and the nickel-copper bob swings as a pendulum. It loses enough heat in one or two vibrations to recover its magnetic properties, and it is then again attracted and held by the magnet. The operation is repeated as soon as the nickel is again heated by the lamp.

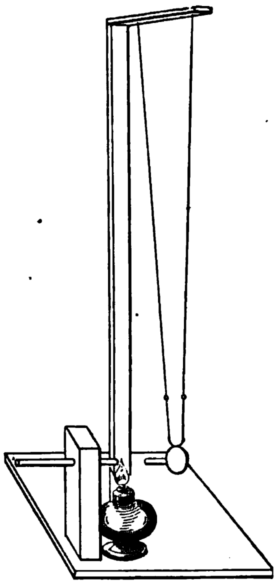


Fig. 290

459. Unit Magnetic Pole. — A *unit magnetic pole*, or a magnetic pole of unit strength, *exerts on an equal pole at a distance of one centimeter in air a force of one dyne.* A pole of strength m exerts a force of m dynes on a unit pole in air at a distance of one centimeter. If therefore m and m' are the pole strengths of two magnets, the mutual force between the two poles in air at a distance of one centimeter is mm' dynes. If the poles are of opposite sign, the product mm' is negative, or the negative sign means attraction.

460. Second Law of Magnetic Force. — Toward the end of the eighteenth century Coulomb investigated the law of attraction and repulsion between magnet poles with the following quantitative result:

The force between two magnetic poles in air is proportional to the product of their strengths and inversely proportional to the square of the distance between them.

The magnets must be relatively long and the distance so great that the poles may be regarded as mere points. Combining this law with the definition of unit pole, we may write

$$\mathcal{F} = \pm \frac{mm'}{r^2}. \qquad (81)$$

461. Magnetic Moment. — The *moment of a magnet* is the product of the strength of its poles and the distance between them, or

$$\mathfrak{M} = ml. \qquad (82)$$

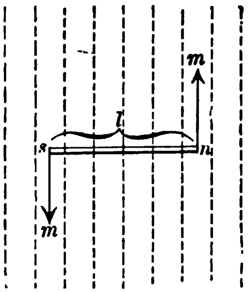


Fig 291

Let the dotted lines of Figure 291 be the direction of the magnetic force in a region where this force is uniform and of unit value, that is, where the force on a unit pole is one dyne; and let *ns* be a magnet with pole strength *m*. Then the force on either pole is *m* and the two forces form a couple. The moment of this couple, when the magnetic axis of *ns* is perpendicular to the direction of the magnetic force acting on the magnet, is *ml*; *ml* is the magnetic moment of the magnet.

II. THE MAGNETIC FIELD

462. Lines of Magnetic Force. — Since a magnet acts on a magnetic needle anywhere in its neighborhood, the space around the magnet is distinguished from other regions by its

magnetic properties. This distinction is expressed by saying that a magnet produces around it a field of force called a *magnetic field.*

If a sheet of cardboard or of thin glass be placed over a bar magnet, and if soft iron filings be evenly sifted over the upper surface, the filings will arrange themselves in curved lines when the paper or glass is gently tapped to set them

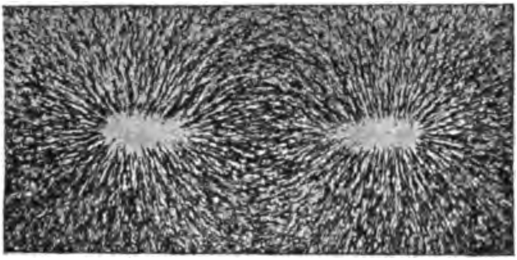


Fig. 292

free. They cling together in lines which diverge from one pole of the magnet and converge again toward the other. These lines are called *lines of magnetic force*, or of *magnetic induction.* The direction of a line of force at any point is the direction of the resultant magnetic force at the point.

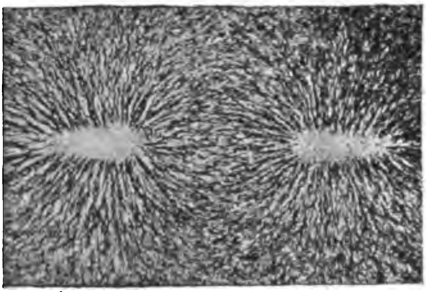


Fig. 293

Figure 292 is reproduced from a photograph made by sifting iron filings on the sensitized side of a photographic plate, with a piece of magnetized watch spring under it.

Figure 293 shows the magnetic induction or field of force between the unlike poles of two similar magnets. The lines from the north pole of the one stretch across to the south pole of the other. Lines of magnetic force are under tension or tend to shorten. They act like stretched cords mutually repelling one another.

The two poles of opposite sign are drawn together by the tension along these lines, or this figure is a picture of magnetic attraction.

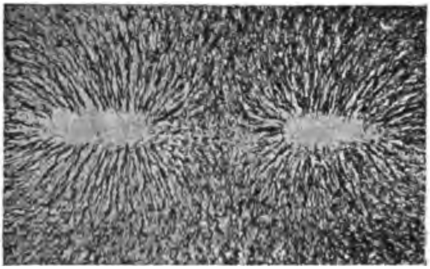


Fig. 294

Figure 294 was made from two like poles. No lines extend across between them. The resiliency of these lines under distortion is such as to force the magnets apart, in order to permit the lines to recover their normal distribution about the poles. This figure is therefore a picture of magnetic repulsion.

463. Direction of Lines of Force. — The direction of a line of force at any point is that of a line drawn tangent to the curve at the point; the direction along it is the same as that in which a north pole is urged. The north pole of a magnetic needle is repelled from the north pole of a bar magnet. Hence, if an observer stands with his back to the north pole of a magnet, he is looking in the direction of the lines of force coming from that pole.

464. Intensity of a Magnetic Field. — The intensity of a magnetic field at any point is the force exerted on unit pole placed at the point. Intensity of field is conventionally denoted by the number of lines of force passing through one square centimeter at right angles to the direction of the field. It is designated by the letter $\mathcal{H}$. The *magnetic flux* through any area s is equal to the product of this area and the strength of field, or $s\mathcal{H}$.

465. Magnetic Flux from Unit Pole. — Imagine a sphere of one centimeter radius described about a unit magnetic pole as a center. The intensity of the field at every point on the surface of this sphere is unity, or one line passes through every square centimeter. Then, since the surface of the sphere is 4π square centimeters, the number of lines belong-

ing to unit pole and passing through the surface of the sphere, that is, the total magnetic flux, is 4π; and for a pole of strength m, the flux is $4\pi m$ lines.

466. Consequent Poles. — A bar of steel or other magnetic body may be magnetized in such a manner that it will have a succession of poles alternating in sign. Thus, in Figure 295 there are north poles at *N N*, and south poles at *S S*. The two poles not at the ends of the bar are *consequent poles*. A consequent pole belongs to two or more magnetic fields. The lines emerging from a consequent N pole go to two or more S poles.

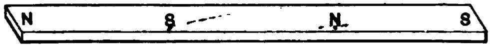


Fig. 295

Consequent poles are commonly used in the field magnets of dynamo-electric machines.

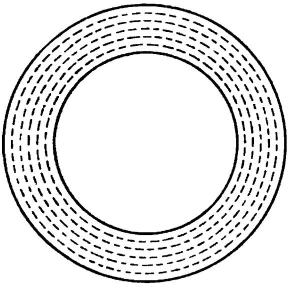


Fig. 296

A ring may be magnetized so as to have consequent poles, or else in such a way that it will not exhibit external magnetic effects. In the latter mode (Fig. 296) there are no poles, that is, no points at which lines of force pass into the air. Such a ring constitutes a closed magnetic circuit, that is, one in which the magnetic flux is entirely in the iron. It has no external magnetic influence, so long as there is no change in its magnetism, because there are no external lines of force. Closed magnetic circuits are more retentive of magnetism than open ones.

III. THEORY OF MAGNETISM

467. Magnetization and Mechanical Stress. — Joule discovered that an iron rod increases in length when it is magnetized longitudinally. He concluded that if it were magnetized circularly, so that the lines of magnetization are circles

around the axis of the rod, it should shorten. This conclusion he verified by experiment.

Bidwell extended Joule's discoveries by showing that at a certain degree of magnetization the elongation reaches a maximum, and that for magnetizing forces beyond that point the elongation is less and less, until finally the dimensions of the iron are unchanged; any increase of the magnetizing force beyond this latter point causes the rod to shorten. Effects of the same kind occur in rings forming closed magnetic circuits; the diameter is increased by small magnetizing forces and is decreased with larger ones.

A circularly magnetized iron wire, when twisted, becomes magnetized longitudinally; and, conversely, torsion in weak fields diminishes longitudinal magnetization and produces circular magnetization. It is conversely probable that the superposition of circular and longitudinal magnetizations will result in torsional strain. Wiedemann demonstrated that this is true in the case of iron. With small magnetizing forces the twist is in one direction, but when the magnetizing forces are large there is a reversal in the direction of the twist.

468. Magnetism Molecular. — A great many facts point to the conclusion that magnetism is a property belonging to the smallest particles composing a body. If a piece of magnetized watch spring be broken in halves, each half will be a magnet with its poles pointing in the same direction as in the original magnet. Smaller subdivision of the spring simply increases the number of poles without destroying the magnetism. The inference is that the ultimate particles or molecules of iron and steel are magnets, and that they are naturally and permanently such.

If a glass tube be nearly filled with fine iron filings, it may be magnetized by stroking it from one end to the other with one pole of a strong magnet. If it then be shaken so as to rearrange the particles, all signs of magnetism disappear. The demagnetization of an iron bar by strong vibration is a similar phenomenon.

Beetz deposited iron electrolytically in a thin line on silver parallel to the lines of induction in a strong magnetic field. The iron was found to be so highly magnetized that no more permanent magnetism could be induced in it.

469. Ewing's Theory of Magnetism. — According to Weber the molecules of iron and other paramagnetic bodies are natural magnets; but in the unmagnetized state of the mass, their magnetic axes lie in all directions crisscross. Ewing has shown that the more probable arrangement of the particles is in closed magnetic circuits, or perhaps stable configurations, under the action of their mutual magnetic forces. A group of such molecules arrange themselves so as to satisfy their relative attractions and repulsions. To illustrate his theory Ewing constructed a model, consisting of lozenge-shaped magnets pivoted on points and arranged at equal distances in a horizontal plane. Any small number of these may group themselves into several stable configurations. After agitation they settle down into groups of equilibrium. With a small external magnetizing force these needles turn through a small angle only; when the force reaches a larger value, some of the needles suddenly turn around and new groupings result, with most of the needles pointing in the direction of the magnetizing force. Any further increase of the magnetizing force produces but little additional effect. These three stages correspond to three similar ones often observed in the magnetization of iron by electric currents (§ 602).

IV. TERRESTRIAL MAGNETISM

470. The Earth a Magnet. —The behavior of a magnetic needle in pointing generally not far from north suggested to Dr. William Gilbert, the leading English man of science in the reign of Queen Elizabeth, that the earth itself is a great magnet. Since then overwhelming evidence has accumulated to confirm his suggestion.

Navigators have located the magnetic pole of the northern

hemisphere in extreme North America within the Arctic Circle. In 1831 Sir John Ross found it in latitude 70° N. and longitude 96° 46′ W. In 1907 Amundsen placed it in latitude 75° 5′ N. and longitude 96° 47′ W. It is not stationary, but is subject to a slow cyclic movement. The magnetic pole of the southern hemisphere has never been reached, but it is probably near latitude 73° S. and longitude 150° E.

Since the N-seeking pole of a magnetic needle points toward the north, it is obvious that the lines of magnetic force about the earth run from the magnetic pole in the southern hemisphere toward the corresponding magnetic pole in the northern hemisphere; in other words, the northern hemisphere of the earth has the polarity of the S-seeking pole of a magnet.

471. Terrestrial Magnetic Induction. — Select a piece of gas pipe 3 or 4 cm. in diameter and about a meter long and carefully free it from magnetism. If it is held horizontally east and west, either end of it will attract both the N-seeking and and the S-seeking poles of a magnetic needle. Gradually tilt it into a vertical position; its lower end will become an N pole and will repel the N pole of the needle. Reverse it and the lower end will again be an N pole and the upper end an S pole. Hold it vertically, or better in the meridian and inclined about 75° below the horizontal toward the north, and strike it a sharp blow on the upper end with a hammer. It has now acquired permanent magnetism with the N pole at the lower end. By reversing it and striking it on the other end the polarity may be reversed, and by graduating the strength of the blow the pipe may be nearly or quite demagnetized.

The earth as a magnet acts inductively on the pipe, as any other magnet does on a piece of iron, putting it under magnetic stress. The vibration due to the blow gives a certain freedom of motion to the molecules, and they arrange themselves to some slight extent under the influence of the earth's

magnetic force. When the molecules are thus arranged, the pipe is a magnet.

Bars of iron or steel in a vertical or in a horizontal north-and-south position acquire magnetism by induction from the earth. This is particularly true if they are subjected to frequent jarring or vibration. Drills, railway iron, beams, and posts are usually found to be magnetized.

472. Magnetic Declination. — The magnetic meridian is the vertical plane coinciding in direction with the earth's magnetic field and containing, therefore, the axis of a suspended magnetic needle. This meridian does not in general coincide with the geographical meridian. The angle between the two is called the *magnetic declination.* The declination is east or west according as the N pole of the needle points to the east or to the west of the geographical meridian. The existence of magnetic declination was not known in Europe until the thirteenth century and was first distinctly shown on a map in 1436. To Columbus belongs the undisputed discovery that the declination is different at different points on the earth's surface. In 1492 he discovered a place of no declination in the Atlantic Ocean north of the Azores.

Lines connecting points of equal declination are called *isogonic lines;* a line of no declination, separating regions in which the declination is westerly from those in which it is easterly, is called an *agonic line.* The magnetic needle everywhere on this line points due north. Such a line in 1900 ran from the north magnetic pole across the eastern end of Lake Superior, through Michigan, Ohio, West Virginia, and South Carolina, and left the mainland near Charleston on its way to the magnetic pole in the southern hemisphere. The magnetic declination at Augusta, Maine, is about 16° W. and at Tacoma, Washington, 23° E.

473. Magnetic Inclination or Dip. — If a magnetic needle be carefully balanced on a horizontal axis through its center of gravity before it is magnetized, its N-seeking pole after

magnetization will incline below the horizontal in the northern hemisphere by an angle ranging from 0° to 90°. This angle is called the ***inclination*** or ***dip***. Norman, a London instrument maker who first measured the dip in 1576, made a ***dipping needle***, which is a magnetic needle free to turn about a horizontal axis in a vertical plane and provided with a graduated vertical circle (Fig. 297). The magnetic poles of the earth are points where the dip is 90°; the dip at the magnetic equator is 0°. The dip in London for 1900 was 67° 9′ and in Washington 70° 18′. It reached its maximum value in London in 1820 and has since been decreasing. Lines of equal inclination on the earth's surface are called ***isoclinic lines***. The ***magnetic equator*** is a line of no inclination in the vicinity of the geographic equator.

Fig. 297

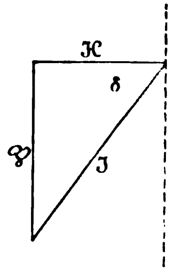


Fig. 298

474. Magnetic Intensity. — The intensity or strength of the earth's magnetic field may be measured in terms of the force in dynes acting on unit magnetic pole. The actual measurements made are those of the angle of dip and the horizontal component of the intensity. If δ is the angle of dip (Fig. 298), then obviously the following relations hold:

$$\frac{\mathcal{H}}{\mathcal{T}} = \cos\delta, \quad \frac{\mathcal{V}}{\mathcal{T}} = \sin\delta, \quad \text{and} \quad \frac{\mathcal{V}}{\mathcal{H}} = \tan\delta.$$

The value of $\mathcal{T}$ at Washington is about 0.6 dyne, and that of $\mathcal{H}$ is about 0.2 dyne per unit pole.

475. Magnetic Variations. — Several distinct changes or variations in the magnetic elements of the earth are constantly taking place. Among them are the following:

1. *Diurnal Variation.* The declination of the magnetic needle has a daily period. During the morning hours the swing of the N-seeking end is toward the east and it reaches its most easterly elongation in general before 9 A.M. It then pauses, turns westward, crosses its mean position, and by 3 P.M. reaches its most westerly elongation. Some time before midnight it again crosses its mean position on the eastward swing, and then repeats the cycle of the day before. The amplitude of its oscillation is about 14′ in the northern portion of the United States and about 4′ in the most southern. These limits vary with the season, being greater in summer and less in winter.

2. *Secular Variation.* In adding to the daily oscillations of the needle, the mean position about which it oscillates is subject to a slow change of long period. This secular variation is different for different localities and for different periods of time. The change in the value of the declination has been recorded in London for more than 300 years. In 1660 it was zero, and it attained its maximum westerly value of 24° in 1810. In 1900 it had decreased again to 15°. If this secular change is periodic, it has a period of about 470 years. The annual change on the Pacific coast is about 4′, and in New England about 3′.

3. *Extraordinary Variations.* In addition to the regular diurnal and slow secular variations, magnetic recording instruments show sudden changes in the magnetic elements, due to some violent terrestrial or cosmic events. For example, such instruments usually furnish a record of earthquake shocks corresponding very nearly in time with the record obtained from the seismograph. While the connection between earthquake shocks and magnetic disturbances has been clearly established, the reason for it has not yet been made out.

476. Terrestrial Magnetism and Sun Spots. — The occurrence of large sun spots and solar outbursts has often been followed by marked magnetic disturbances, as shown by the automatic records of the magnetograph. Recently Professor Hale has obtained evidence at the Mount Wilson Solar Observatory that a sun spot is a magnetic field, or has polarity. The evidence cannot be reviewed here, but it is conclusive in respect to polarity, sometimes N-seeking and sometimes S-seeking; for the effect on spectroscopic lines is qualitatively the same as one obtains in the laboratory with the source of light in a magnetic field (§ 661). The sudden production of these solar magnetic poles during the initial stages of a solar vortex or whirlpool does not appear to affect the earth's magnetism by the reaction between the solar and terrestrial magnetic fields.

CHAPTER XVII

ELECTROSTATICS

I. ELECTRIC CHARGES

477. Electricity and Electrification. — The simple fact that a piece of amber (a fossil gum), rubbed with a flannel cloth, acquires the property of attracting bits of paper, pith, or other light bodies, has been known since about 600 B.C. But it appears not to have been known for the following 2200 years that any bodies except amber and jet were capable of this kind of excitation. About the year 1600 Dr. Gilbert discovered that a large number of substances possess the same property. These he styled *electrics* (from the Greek word for amber, *electron*), but the word *electricity* to designate the invisible agent to which the phenomenon should be referred appears to have been introduced by Boyle in 1675.

A body excited in this manner is said to be *electrified; electrification* is the result of the work done in electrifying, or charging with electricity. Electrification, or electricity under pressure, is a form of potential energy, just as air under pressure or water at an elevation above the earth represents potential energy. Electricity, air, and water are not themselves energy, but only the agents which store or carry energy.

Electricity appears to be as indestructible as matter and energy. Its distribution is subject to control. It may represent energy of stress or energy of motion; but when its energy has been spent in producing physical effects, its quantity remains unchanged. None can be created or generated and none destroyed.

478. Electrical Repulsion. — If several pith balls are suspended by fine linen threads from a glass rod, and an electrified glass tube is brought near them, they are first attracted to the tube but soon fly away from it and from one another (Fig. 299). Their mutual repulsion continues after the tube is withdrawn; and, if the hand is brought near them, they move toward it as if attracted, showing that the balls are electrified.

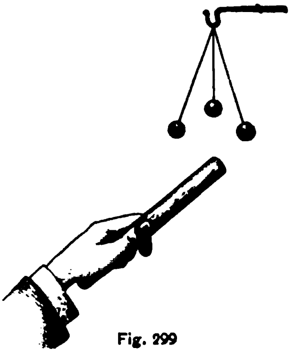


Fig. 299

It thus appears that bodies are electrified by coming in contact with other electrified bodies, and that electrification is shown by repulsion as well as by attraction.

479. Attraction Mutual. — Boyle discovered that the attraction between an electrified and an unelectrified body is mutual. Excite a glass tube by rubbing it with silk, and lay it in a light stirrup suspended by a silk thread (Fig. 300). If the hand is presented to it, it will swing around by the attraction. Force, whatever its origin, is of the nature of a stress, and action and reaction are equal.

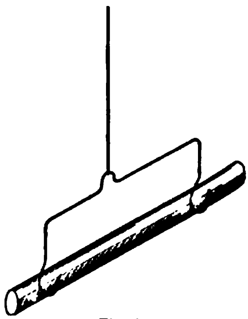


Fig. 300

480. Two Kinds of Electrification. — Not all electrified bodies repel one another. If a second excited glass tube be presented to the one hung in the stirrup (Fig. 300), there will be mutual repulsion between them. On the other hand, a rod of gutta-percha or a stick of sealing wax excited by rubbing with flannel will attract the electrified glass tube; also, if a pith ball be charged by contact with the rubbed sealing wax, it will be repelled

by the sealing wax, but attracted by the glass tube rubbed with silk.

From such facts as these, Du Fay drew the inference that there are *two kinds of electrification.* Electricity manifests itself under two aspects analogous to the two poles of a magnet. Du Fay called the two kinds *vitreous* and *resinous* respectively; but since the result of the friction depends on the rubber as well as on the material rubbed, the terms *positive* and *negative* introduced by Franklin have been adopted everywhere as preferable. The kind of electrification which makes its appearance on glass when rubbed with silk is called *positive*, while the kind excited by rubbing sealing wax with flannel is called *negative.*

These terms are purely arbitrary, and if their choice had been postponed until the present, it is highly probable that many phenomena of recent discovery would have dictated the exchange of the words positive and negative as applied to electrification.

481. First Law of Electrostatics. — All the phenomena similar to those described in the last article are included under the first law of electrostatics, namely: —

Like kinds of electrification produce mutual repulsion between the bodies charged; unlike electrifications produce mutual attraction.

482. Conductors and Nonconductors. — To all substances which do not show electrification by friction Gilbert gave the name "nonelectrics"; but in 1729 Stephen Gray discovered that these substances convey away the "electric virtue" as fast as it is excited. If a metal rod be held by a dry glass handle, it can be electrified by rubbing with silk. Gray conveyed electric charges to a distance of 700 feet by means of a hempen thread supported by silken loops.

Ever since Gray's discovery, substances have been classified as *conductors* and *nonconductors* or *insulators.* These

are only relative terms, for all substances may be arranged in a graded series with the best conductors at one end and the poorest at the other. None conduct perfectly and none insulate perfectly. The difference in relative conductivity, however, is enormous. Thus, silver conducts more than 60 times as well as mercury, and 2500 times as well as gas carbon, while the conductivity of mercury is 100,000 times greater than that of dilute sulphuric acid.

483. The Electroscope and the Electrometer. — The electroscope is an instrument for detecting electric charges. The most common form is the gold leaf electroscope. Through the top of a glass jar passes a brass rod, terminating in a ball, or horizontal plate above, and in two strips of gold leaf on the inside hanging parallel and close together. The metal rod must be insulated from the glass either by a heavy coat of melted shellac or by a stopper of sulphur.

If the knob of an electroscope be touched with an excited glass tube, the gold leaves will repel each other with positive charges. The approach of any other charged body will cause them to diverge more widely if the body is charged positively, and to approach each other if it is charged negatively.

The study of ionization and of radioactivity in recent years has led to a modification of the electroscope with the object of making it a measuring instrument. Of the many forms proposed, that of Wilson is typical. The indicating system consists of a rigid piece of flat brass B (Fig. 301), to which is attached a narrow strip of gold leaf G. This system is supported by a block of sulphur I, which in turn is suspended by a rod fitting tightly in a block of ebonite E. A charging wire W passes through the ebonite and is bent at right angles at the bottom. By rotating the upper bent end of W, the arm at the bottom may be brought in contact with the brass strip; it may be disconnected as soon as the system is charged. In some instruments the gold leaf is viewed with a microscope of low magnifying power.

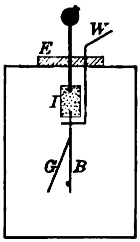


Fig. 301

For measuring small electrostatic potential differences (§ 500) Lord

Kelvin devised the quadrant electrometer. The essential parts are the cage, or quadrants, and the needle. The four insulated quadrants together form a short hollow cylinder with parallel ends (Fig. 302). The needle, a long, thin piece of aluminum with broad rounded ends, is suspended by a fine wire or fiber so as to turn in a horizontal plane inside the four quadrants. It is connected at the bottom with the jar *B* by a fine platinum wire dipping into sulphuric acid. Opposite quadrants are connected electrically.

Fig. 302

If all the quadrants are equally charged, the needle will take a position depending only on the torsion of the suspending fiber; but if two of the diagonally opposite quadrants are charged positively and the other two negatively, the needle being positive, the latter will rotate in one direction or the other, according to the connections, until the couple due to the electrostatic attractions and repulsions is just balanced by the torsion of the suspension.

484. Positive and Negative Electrifications Equal. — When a body is electrified by friction, the body rubbed and the rubber are equally electrified, but with charges of opposite sign. The equality consists in the ability of the one charge to exactly neutralize the other.

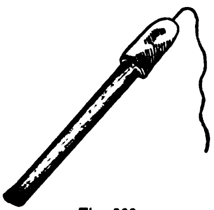


Fig. 303

If a stick of sealing wax, provided with a flannel cap with a silk cord attached (Fig. 303), be excited by turning it around a few times inside the cap, it will show no signs of electrification if presented to the knob of an electroscope without removing the cap; but if the cap be withdrawn by the cord, the sealing wax will cause the gold leaves to diverge less and the flannel cap will cause them to diverge more, provided the electroscope is charged positively.

The inferences are: first, *that one kind of electrification is not produced without the other;* second, *that the two kinds are produced in equal quantities.* The electrification of a body consists in the separation of two equal charges of opposite sign against their mutual attraction. It follows that the medium between them is strained by the operation, and work is done. The slightest positive charge at one point always means an equal charge of the opposite sign as near it as the conductance of the insulators separating the two charges permits.

485. Charge External. — When a conductor is electrified by friction or by a charge conveyed to it from some external source, the charge always resides on the outside. A simple demonstration of this principle may be made by means of a hollow metal sphere with a hole at the top and insulated on a glass stem as a support (Fig. 304). It may be tested by means of a proof plane (Fig. 305), which consists of a small metal disk cemented to one end of an ebonite or shellac handle. If the proof plane be applied to the outside of the charged sphere, a small charge may be removed and tested by an electroscope. If the proof plane be passed through the hole in the sphere and touched to the inner surface, it will not show any trace of electrification. If in fact it be charged from the outside of the sphere, and then be made to touch the interior, it will lose all its charge and will show none on withdrawal.

Fig. 304

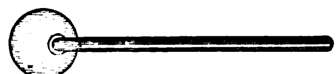


Fig. 305

486. Distribution of the Charge. — The quantity of electricity on a square centimeter of the surface of a conductor, or the ratio of the quantity on any small area to the area itself, is called the *surface density*. The distribution of an electric charge on an insulated conductor is not such as to give uniform surface density over it, except in the case of a sphere remote from other conductors and electrified bodies. The surface density on different parts of a conductor may be examined by means of a proof plane and an electroscope. After the proof plane has been brought in contact with the rounded end of a charged cylindrical conductor, it will produce a much greater divergence of the leaves of an electroscope than after it has touched the side of the cylinder.

On a cylinder with rounded ends, the surface density is greatest at the ends.

On a flat disk the density is much greater at the edges than on the flat surface; the distribution over the latter is fairly uniform except near the edges.

The surface density is greatest on those parts of a conductor which project most and have the greatest curvature outwards. On sharp points, such as that of a needle, the density is very great, and the charge escapes rapidly from them into the air. It is for this reason that the edges of conductors are rounded and made smooth.

487. Second Law of Electrostatics. — The law of the force between two electric charges was first investigated by Coulomb. He demonstrated that *the stress between two electric charges is directly proportional to the product of the two quantities, and inversely proportional to the square of the distance between them.*

The law of distance does not hold unless the charged conductors are very small in comparison with the distance between them. The distribution of the charges on the two bodies is then not appreciably affected by their mutual action.

The second law of electrostatics may be succinctly expressed in terms of algebraic symbols as follows: Let q and q' denote the electric charges of the two small bodies, d the distance between them *in air*, and C a proportionality factor. Then Coulomb's results are expressed by the equation

$$F = C\frac{qq'}{d^2}. \tag{83}$$

488. Unit Charge. — The definition of unit charge or unit quantity is so chosen that the constant C in the foregoing equation becomes unity. *The electrostatic unit charge is that quantity which exerts on an equal quantity one centimeter distant in air a force of one dyne.*

It is necessary to say "in air" because, as will be seen later, the force between two charged bodies depends on the nature of the medium between them.

489. Nature of Electrification. — It was suggested by Faraday, and a great many facts tend to confirm his view, that the electrification of a body produces a strained condition of the ether around it. Conductors differ from insulators in this respect: in the former the molecular mobility is such that the state of strain is continually giving way; while in the latter considerable distortion is possible before the molecular structure yields to the stress. The phenomena of attraction and repulsion exhibited by electrified bodies are due to the effort of the strained ether in and about the bodies to return to its normal condition. In producing electrification, work is done in distorting the medium; hence electrification is a form of potential energy.

490. Electric Field. — Any region within which the medium is under stress is a *field of force;* and an *electric field* is one in which the stress is due to an electric charge. The intensity of an electric field at any point is defined as the force sustained by a unit charge placed at the point. A unit electric field is one in which unit charge is acted on with a force

of one dyne. Electric intensity is measured in the same fundamental units as mechanical forces.

491. Lines of Electric Force. — An electric field, like a magnetic field, is conveniently represented by *lines of force*. At any point in the field a line of force takes the direction of the electric force at that point. It is conceived to start

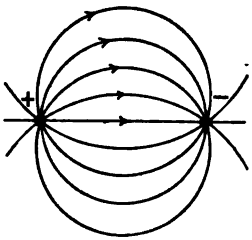


Fig. 306

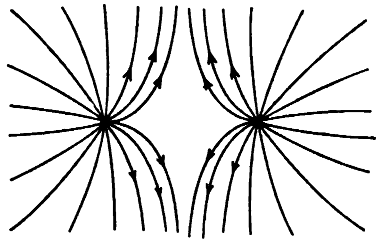


Fig. 307

on a positive charge and to end on a negative one. The positive direction is therefore the direction in which a small positively charged body tends to move along the line of force.

For a positive charge on a very small body, the lines of force are radial lines emanating from the charge. For two equal charges of opposite sign on very small bodies, the lines of force run as in Figure 306. For like charges, they take the form shown in Figure 307. No two lines cross, for if they did, a charge placed at the intersecting point would tend to move in two directions at the same time.

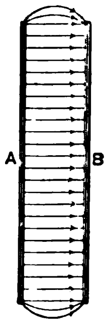


Fig. 308

The lines of force in an electric field tend to shorten like stretched cords. When one electrified body attracts another, they are drawn together by these taut lines extending from one to the other. If two parallel plates face each other and are charged oppositely (Fig. 308), the lines of force stretch across from the positive to the negative, and the tension in the medium pulls the plates together. When the

plates are close together, the field is nearly uniform and the lines of force are parallel except near the edges.

II. ELECTROSTATIC INDUCTION

492. Induction Phenomena. — A charged conductor exerts influence, or acts inductively, on neighboring bodies. If it be charged positively, lines of electric force spring from it and proceed to an equal negative quantity on adjacent bodies; the influence or induction is exerted along these force lines.

Let an insulated sphere *A* (Fig. 309), charged positively, be placed near an insulated cylindrical conductor *B*. If *A* and both ends of *B* be now examined by means of a proof plane and an electroscope, it will be found that the charge on *A* has been redistributed, so that the surface density on the side toward *B* is greater than on the remote side; also, that the end *a* of the cylinder has negative electrification, the central portion is neutral, and the end *b* has positive electrification. The surface density at *b* is less than at *a*, and the neutral line is somewhat nearer *a* than *b*.

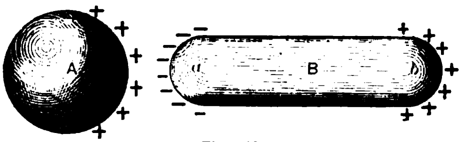


Fig. 309

As soon as *A* is removed or is discharged by connecting with the earth, all signs of electrification on *B* disappear. The electrification of *B* is due to the charge on *A*, but is not at the expense of that charge. The separation of the positive and negative charges on *B* through the influence of the charge on *A* is called *electrostatic induction*, or electrification by *influence*.

493. Charging by Induction. — The next step is to give to *B* a charge of one sign by induction. While it is still under the influence of the charge on *A*, let it be connected for an instant with the earth. The effect is to enlarge indefinitely

the dimensions of *B*, and the positive electrification goes to the remote part of this enlarged conductor, that is, to the earth. If now *A* be removed, while *B* remains insulated, the negative charge on the latter will distribute itself over the entire conductor, and *B* is said to have been *charged by induction.* This induced charge is of the opposite sign to the inducing charge on *A*.

Charging by induction may also be referred to lines of force in the enveloping medium. When *B* (Fig. 309) is brought into the field of *A*, some of the lines of force springing from *A* terminate on the nearer end of *B*; an equal number leave *B* at the distant end; that is, the end *a* is charged negatively and *b* positively. When *B* is connected to the earth by a conductor, the stress in the medium, represented by the lines running from *B*, is removed, and only that between *A* and *B* remains. Hence, *B* has then a negative charge only.

494. Electrification with Like Charges by Induction. — It is quite possible to charge by induction so that the induced charge shall be of the same sign as the inducing charge. Imagine the conductor *B* provided with sharp points at the end *a* (Fig. 310), and a circular glass platé revolving with its edge between *A* and *B*. The negative charge on *a* will then acquire so great a density on the points that they will discharge it on the revolving plate. If another

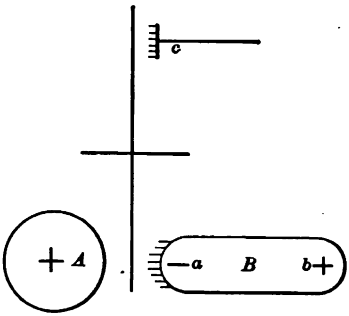


Fig 310

row of points *c*, connected with the earth, be placed opposite the same side of the glass plate, but away from the inductive action of *A*, the revolving plate will give up to *c* the negative charge acquired at *a*, and *c* will convey it to the earth. In

this way B is left with a positive charge. The electrification still represents energy, for work is done in turning the glass plate against the attraction of the unlike charges on it and A.

495. Attraction due to Induction. — The simple facts of electrostatic induction furnish an explanation of the attraction between electrified and unelectrified bodies. The induced charge of opposite sign always accumulates on the part of the conductor nearest the inducing charge, while the repelled charge goes to the most distant parts of the conductor, or to the earth if a conducting path is furnished.

If an excited glass rod C (Fig. 311) be presented to an uncharged pith ball suspended by a silk thread, a negative charge will be induced on the ball at a and positive at b. Since the former is nearer C than the latter, the attraction will be greater than the repulsion, and the pith ball on the whole will be attracted. If it be touched while under induction, the repelled positive charge will disappear and the attraction will be increased.

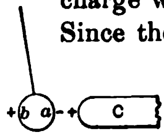


Fig. 311

496. The Inducing and Induced Charges Equal. — The charge induced on a conductor can never exceed the inducing charge. If all the force lines from the inducing charge go to the induced charge on the conductor, the two charges are equal. If, for example, a charged ball be nearly surrounded by a hollow conductor (Fig. 312), all the lines of force from the ball A end in the induced charge on the inclosing conductor. No sensible number escape through the small opening. A negative charge then spreads over the interior of B equal in quantity to the positive charge on A.

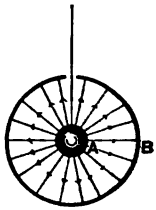


Fig. 312

In this case the charge on B is on the inside instead of the outside. If B is insulated while under the inductive influence of A, and A is then removed without making contact with B, the negative charge on B will all go to the outside.

497. Faraday's Ice Pail Experiment. — Faraday employed a pewter ice pail as a convenient hollow conductor to test the question of equality between the induced and the inducing charges. *A* is a section of a well-insulated pail (Fig. 313). The outside is connected to a gold leaf electroscope *E*. A positively charged ball *C* is let down into the pail by means of a silk thread. As soon as it enters the pail the gold leaves begin to diverge, and the divergence increases until the ball reaches a certain depth, depending on the relative dimensions of the pail. Beyond this depth the divergence remains constant. The divergence increases up to the point where all the lines of force from *C* terminate in the negative charge on the inside of the pail.

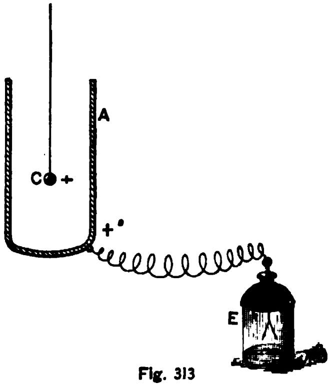


Fig. 313

If now the ball be allowed to touch the pail, not the slightest change in the divergence of the gold leaves can be detected. The inference is that the positive charge on the outside of the pail, when the ball is acting inductively on it, is exactly equal to the charge communicated to the pail when the ball makes contact.

The experiment was varied by touching the pail while under the influence of the charge on the ball. The gold leaves collapsed. But when the ball was withdrawn, they again diverged to the same extent as before, but with a negative charge. If next the charged ball was replaced and made to touch the pail, all signs of electrification disappeared, indicating again an equality between the induced charge and the positive conveyed by the ball.

498. The Electrophorus. — The simplest induction machine is the *electrophorus* invented by Volta. By means of it an

indefinite number of small charges may be obtained by induction from a single charge due to friction. It consists of a plate of hard rubber or resin and a metal cover provided with an insulating handle (Fig. 314). The hard rubber plate usually rests on a metal plate or sole.

Fig. 314

To use the electrophorus, the hard rubber or resin is electrified by beating it with a catskin, and the metal cover is placed on it. Since the cover touches the nonconducting disk in a few points only, it does not remove the negative charge due to the friction of the catskin, but it is acted on inductively by the negative charge on the disk. The cover is then touched for an instant with the finger; the repelled negative charge passes off, leaving the cover with an induced positive charge on its lower surface. The cover may then be lifted by the insulating handle; the bound charge becomes free, and is available for charging other bodies.

Evidently this process may be repeated an indefinite number of times without removing any appreciable part of the charge from the hard rubber plate. It is, however, slowly dissipated in damp air or if the hard rubber is not dry.

III. ELECTRIC POTENTIAL

499. Definition of Electric Potential. — The term *potential*, introduced by Green in England in 1828, was originally a mathematical function, but it has now the greatest practical significance in the science of electricity.

First. *Mutual potential energy.* Consider two unlike electric charges; the *mutual potential energy* of such a system in any given position is the work which must be done

against their mutual attraction in separating them to an infinite distance, or in conveying one to the boundary of the field produced by the other. It should be carefully noted that the force worked against is purely electrical; all others are excluded.

Second. *Potential.* The *potential* at any point, due to a given charge, is the mutual potential energy of this charge and *unit quantity* of electricity placed at the point. In other words, it is the work which must be done on a unit charge in carrying it from the point to an infinite distance, or to the boundary line of the field of force due to the given charge.

For convenience the potential of the earth is usually taken to be the arbitrary zero of potential, just as the sea level is an arbitrary level from which altitudes are measured.

500. Difference of Potential.—Consider two points A and B, and let the potentials at these points be denoted by V_1 and V_2 respectively. Since work equal to V_1 ergs is required to convey unit charge from A to infinite distance, and work V_2 ergs from B to an infinite distance, it is obvious that the work done against the electric forces in displacing a positive unit charge from one point to the other is $V_1 - V_2$. This work is independent of the path followed in going from A to B; otherwise it would be possible, by making a quantity of electricity circulate between A and B by suitable paths, to gain energy without a corresponding expenditure of work. If the work expended in conveying the charge from A to B by one path should be less than the energy recovered by allowing it to return by electric forces from B to A, energy would be accumulated in every cycle, and this is contrary to the law of conservation of energy. The same principle applies to work done against gravity between two points at different levels. The work done is independent of the path traversed between the two points; otherwise, "perpetual motion" and continuous work without a corresponding expenditure of energy would be possible.

501. Equipotential Surfaces.—An equipotential surface is the analogue of a level surface. The potential at all points of an equipotential surface is the same. There is then no difference of potential between points on an equipotential surface, and no work is done in carrying a charge of electricity between any two points on the same equipotential surface.

The surface of a charged conductor is an equipotential surface, for its charge is in equilibrium and there is no force along the surface of the conductor and therefore no potential difference. But a charged conductor produces a field of force, and the only direction in which a charge can be moved without work is at right angles to the lines of force. It follows that the lines of force are everywhere normal to the equipotential surface. They terminate at the charge on the exterior, for there is no force on the inside of a charged conductor.

502. Work between Equipotential Surfaces.—Consider two equipotential surfaces, the potentials of which are V_1 and V_2. The work done in displacing unit charge from the one surface to the other is the difference of potentials, $V_1 - V_2$. It is independent of the path traversed and of the position of the point of departure and the point of arrival on the two surfaces. If a quantity q units is conveyed from one surface to the other, the work done is q times as much as for one unit, or $q\ (V_1 - V_2)$. The numerical measure of the electrical work is thus a product of two factors, one of them a potential difference and the other a quantity of electricity. If the potentials of the two surfaces differ by unity, then one erg of work must be done to convey unit charge from the one surface to the other.

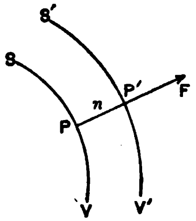


Fig. 315

503. Force in Terms of Potential.—Let S and S' be two equipotential surfaces very near together (Fig. 315), and

let V and V' be their potentials. Let F be the electric force along a normal between P and P'. If n is the distance PP', the work done by the force F in conveying a unit quantity from one of the surfaces to the other is $F \times n$. Then

$$Fn = V - V', \quad \text{and } F = \frac{V - V'}{n}. \tag{84}$$

The electric intensity along a line of force is therefore equal to the decrease of the potential per unit length along the line. In general the intensity in any direction equals the rate of the diminution of potential in that direction. It follows that the intensity between two equipotential surfaces is greatest in the direction in which the distance n is least.

504. Equilibrium of a Conductor. — When a charge of electricity is given to a conductor, it at once distributes itself over the surface and comes to equilibrium. Moreover, since there is no force inside a charged conductor, there is no difference of potential throughout its entire volume, since force equals the rate of variation of potential. All parts of a charged conductor have therefore the same potential.

The surface of an insulated conductor under the influence of a charged one is also an equipotential surface, because the charges on it are in equilibrium and there is no electric flow along it. For example, the potential at a (Fig. 316), due to the positive charge on A, is higher than at the more distant point b; but the negative charge near a lowers the potential of the nearer half of the cylinder, and the positive near b raises the more distant half to the same potential level as at a. If now the cylinder be connected to earth, its potential will be reduced to that of the earth, that is, to zero. The cylinder will remain charged negatively, but its potential will be

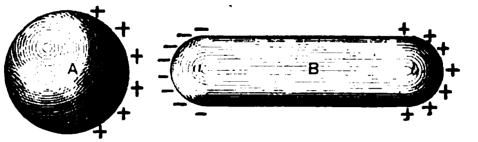


Fig. 316

zero. The positive potential due to the charge on A and the negative due to its own charge everywhere equal each other, and the resultant is zero. It is obvious then that surface density and potential are in no sense the equivalents of each other.

505. Potential equals $\sum\frac{q}{r}$. — The problem is to find the expression for the potential at A, at a distance r from an element q of the charge on the conductor O (Fig. 317). Let B

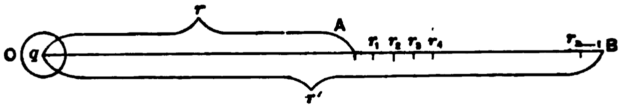


Fig. 317

in the line OA be at a distance r' from O. Let the distance AB be supposed divided into n very small elements, so that the points of division are distant r_1, r_2, r_3, etc., from O.

Then, since the electric intensity is the force on unit charge, the intensity at the distance r is q/r^2, at the distance r_1 it is q/r_1^2, etc. Since r and r_1 are very nearly equal, we may put the equivalent intensity between the two adjacent points r and r_1 equal to q/rr_1. Similar expressions obtain for the other elements of the distance AB.

Hence the work in carrying unit charge

$$\text{from } r \text{ to } r_1 = \frac{q}{rr_1}(r_1 - r) = q\left(\frac{1}{r} - \frac{1}{r_1}\right);$$

$$\text{from } r_1 \text{ to } r_2 = \frac{q}{r_1 r_2}(r_2 - r_1) = q\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

. . . .

$$\text{from } r_{n-1} \text{ to } r' = \frac{q}{r_{n-1}r'}(r' - r_{n-1}) = q\left(\frac{1}{r_{n-1}} - \frac{1}{r'}\right).$$

The whole work done in carrying the unit charge from A to B is the sum of all these small quantities; but in sum-

ming up it will be observed that all the intermediate terms cancel out; hence the work between A and B is

$$q\left(\frac{1}{r}-\frac{1}{r'}\right). \tag{85}$$

This expression is the difference of potential between A and B due to the element of charge q at O.

If we suppose the point B removed to an infinite distance, $1/r'$ becomes zero, and

work from A to an infinite distance is $\frac{q}{r}$.

But by definition this is the potential at A due to the charge q. The expressions for other elements of the charge are similar; and since potential is a scalar quantity, the resulting potential at A is the algebraic sum of the potentials due to the several elements of the charge, or

$$V=\frac{q'}{r'}+\frac{q''}{r''}+\frac{q'''}{r'''}+\ .\ .\ .\ =\sum\frac{q}{r}. \tag{86}$$

506. Potential of a Sphere. — If the sphere has a charge Q, every element q of this charge is at a distance r from the center of the sphere, and the potential at the center due to an element q is q/r, r being the radius of the sphere. The potential at the center due to the entire charge is then

$$\sum\frac{q}{r}=\frac{1}{r}\sum q=\frac{Q}{r}. \tag{87}$$

But as all points of a conductor in equilibrium have the same potential, the potential of every point of the sphere due to a charge Q is Q/r.

A charge uniformly distributed over a sphere acts at external points as if it were collected at its center; hence the potential at any point external to the sphere and distant d units from its center is Q/d.

Problems

1. What would be the potential difference between A and B (Fig. 317) if O were charged with 100 units of electricity, the distance r being 10 cm. and r' 20 cm.?

2. Positive charges, 150, 424, and 300 units, are placed at the three corners A, B, C, of a square 40 cm. on a side. Calculate the potential at the fourth corner D.

3. Positive charges of 50 units each are placed at the three corners of an equilateral triangle whose sides are 100 cm. Find the potential at the center of the circumscribing circle.

4. Find the potential at the center of the square in problem 2, and the work required to carry a unit charge from D to the center.

5. Two charges, 100 positive and 70 negative, are placed 30 cm. apart. Find the potential at a point on the line joining them 70 cm. from the negative charge and 100 cm. from the positive; also, calculate the force on a charge of 40 positive units at the same point.

IV. ELECTRIC CAPACITY AND CONDENSERS

507. Definition of Capacity. — The electric capacity of a conductor is defined as the *ratio of its charge to its potential when all other conductors within its field are at zero potential.* This definition is equivalent to saying that the numerical value of its capacity is the number of units of electricity that will raise its potential from zero to unity. In symbols, if C denotes capacity,

$$C = \frac{Q}{V}. \tag{88}$$

Also $$Q = CV \text{ and } V = \frac{Q}{C}.$$

508. Capacity of an Insulated Sphere. — The capacity of a sphere in air remote from other conductors is numerically equal to its radius in centimeters. For the potential of such a sphere is Q/r. Hence

$$C = \frac{Q}{V} = Q \div \frac{Q}{r} = r. \tag{89}$$

The radius is in centimeters because the centimeter is the unit of length employed in defining the unit of charge.

509. Condensers. — Two conductors placed near each other with an insulator, called the *dielectric*, between them form with the dielectric a *condenser*. The effect of the additional conductor and the dielectric is to increase the charge without any increase of potential. In other words, the capacity of the one conductor is greatly increased by the presence of the other. The ratio of the charge on either conductor to the potential difference between two is the *capacity of the condenser*.

510. Capacity of Two Concentric Spheres. — It is not difficult to calculate the capacity of a condenser when the plates have certain simple geometric forms. Let r be the radius of the inner sphere and r' that of the outer one (Fig. 318), and assume that the outer surface is connected to the earth, so that its potential is zero. Then all the lines of force from the insulated charged sphere A run to the outer sphere B, and their charges are equal and of opposite sign, $+Q$ and $-Q$.

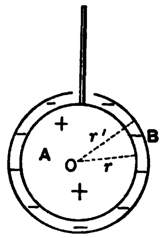


Fig. 318

The potential at O, the common center of the two spheres, is

$$V = \frac{Q}{r} - \frac{Q}{r'} = Q\left(\frac{1}{r} - \frac{1}{r'}\right).$$

This is the potential of the entire inner sphere because the potential inside a charged conductor is the same as at any point on its surface. From the last equation

$$\frac{Q}{V} = \frac{rr'}{r' - r} = C \text{ (§ 507)}.$$

When $r' - r = t$ is very small, that is, when the two spherical surfaces are very near each other, the capacity becomes very large. The expression for the capacity is then

$$C = \frac{rr'}{t} = \frac{r(r+t)}{t},$$

where t is the thickness of the dielectric, which is here assumed to be air. When t is very small compared with r, the expression for the capacity becomes

$$C = \frac{r^2}{t} = \frac{4\,\pi r^2}{4\,\pi t} = \frac{A}{4\,\pi t}, \tag{90}$$

where A is the surface area of the inner sphere.

511. The Leyden Jar. — The Leyden jar was the earliest known form of electric condenser. It derives its name from the fact of its accidental discovery in the city of Leyden. As now made, it consists of a wide-mouthed jar of thin flint glass, coated inside and out with tin foil to about three fourths its height (Fig. 319). The metal knob is connected to the inner coating by a rod terminating in a piece of chain.

Fig. 319

The jar may be charged by holding it in the hand, touching the knob to one electrode of an influence machine, and bringing the outer coating so near the other electrode that a series of sparks pass across. If charged to too high a potential, it will discharge along the glass over the top. It may be safely discharged by a discharger (Fig. 320) held by the glass handles, one ball being brought into contact with the outer coating and the other with the knob.

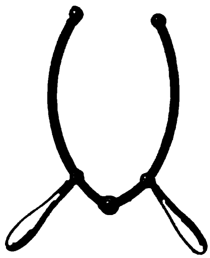


Fig. 320

If A is the area of the tin foil and t the thickness of the glass, with air as the dielectric the capacity is approximately the same as with concentric spheres of the same conducting area A, that is,

$$C = \frac{A}{4\,\pi t}.$$

The effect of interposing glass instead of air between the two coatings is to increase the capacity by a factor K, so that

$$C = K\frac{A}{4\pi t}. \tag{91}$$

K is a constant of the dielectric depending on the kind of glass, and varying in value from about 3 to 7 for different specimens.

512. Residual Charge. — If a Leyden jar be left standing for a few minutes after it has been discharged, the two coatings gradually acquire a small potential difference and a second small discharge may be obtained from it, due to the *residual charge*. Several of them, of decreasing intensities, may sometimes be observed. The potential of the residual charge depends on the potential difference to which the jar has been charged, the length of time it is left charged, and the kind of glass forming the dielectric.

513. Seat of the Charge. — The Leyden jar with removable coatings (Fig. 321) is due to Franklin. By means of it he demonstrated that the charge is apparently on the surface of the glass. The real significance of the experiment is somewhat different.

Fig. 321

The inner and outer coatings are metallic and fit the glass jar. If the jar be charged in the usual manner and be placed on an insulating stand, the inner metal vessel may be lifted out by means of the curved rod; the glass jar may then be withdrawn from the outer vessel. The two metallic coatings are now completely discharged; but after the parts are all replaced, the jar may be discharged with a bright flash.

The electrification of the jar is to some extent a phenomenon of the glass. During the charging of the jar the glass

is strained; the conductors carry the charge and facilitate the release from strain. Thin glass jars may be crushed or perforated by overcharging. The residual charge indicates that the glass acts as if it were distorted, like a twisted glass fiber, and that it does not return at once to its normal unstrained state when the jar is discharged.

514. Energy of a Charged Conductor. — The energy expended in carrying Q units of electricity through a potential difference $V - V_0$ is $Q(V - V_0)$. If the charge Q were transferred from the earth, whose potential is zero, to a conductor whose potential remained V units, the work done would be QV. But in charging a condenser, or any other conductor, the potential is zero at the beginning of the charging process and V at the end. Since the potential at every instant is proportional to the charge, the mean potential to which the charge is raised is $\frac{1}{2}V$, and the work done in charging the condenser, which equals the energy of the charge, is

$$W = \tfrac{1}{2}QV.$$

Since $Q = CV$, other equivalent expressions for the energy are

$$W = \tfrac{1}{2}CV^2 = \tfrac{1}{2}\frac{Q^2}{C}. \tag{92}$$

Illustrations of similar expressions for mechanical work done are readily found. Thus, the work done against gravity in building a tower of uniform cross section (corresponding to capacity) is equal to the product of the mass of the material and half the height, or $\frac{1}{2}MH$. To build the tower twice as high requires four times as much work, because double the mass is lifted to twice the mean height.

So also the work done in compressing gas into a cylinder of fixed volume is proportional to half the product of the quantity and the final pressure, or $\frac{1}{2}QP$. To fill it to a pressure twice as great requires forcing in twice the mass of gas against twice the mean pressure, or the work done is proportional to the square of the pressure. Height in the one case and pressure in the other are the analogues of potential.

515. Energy Lost in dividing a Charge. — Let C_1 and C_2 be the capacities of two condensers, and let the first be charged

with a quantity Q. If the potential is V, the energy of the charge is

$$W = \frac{1}{2} C_1 V^2 = \frac{1}{2} \frac{Q^2}{C_1}.$$

After the charge has been divided by connecting the two condensers in parallel, the capacity has been increased to $C_1 + C_2$, and the energy is therefore reduced to

$$W = \frac{1}{2} \frac{Q^2}{C_1 + C_2}.$$

The stored potential energy after the division of the charge is less than before it so long as C_2 has any value in comparison with C_1. If the two capacities are equal, the energy stored after the division is half as great as before it. The other half is represented by the energy of the spark at the moment of the division. Energy is always lost by the division, whatever be the relative values of the two capacities.

516. Energy of Similar Condensers in Parallel. — If n condensers of the same capacity C are charged in parallel, n Leyden jars for example, with all their outside coatings connected together, also their inside coatings (Fig. 322), the capacity of the whole is n times the capacity of a single jar; the effect is simply to increase the size of the coatings.

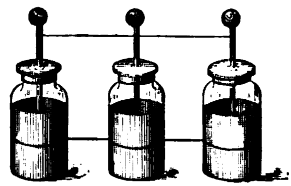


Fig. 322

Since the energy of discharge of a single condenser is $\frac{1}{2}CV^2$, for n condensers of the same capacity it is

$$W = n \times \frac{1}{2} CV^2 = \frac{1}{2} nCV^2.$$

The energy of the charge is proportional to the number of similar condensers.

517. Energy of Condensers in Series. — If several Leyden jars are insulated and the outside of one is joined to the in-

side of the next (Fig. 323), they are said to be connected in series or in "cascade." Let the series be charged to a potential difference of V units between the inside coating of the first and the outside of the last. Then the potential difference between the coatings of each jar is V/n, and the energy of its charge is $\frac{1}{2}\, CV^2/n^2$. The energy of the n charged jars is

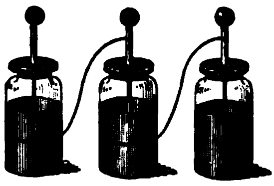


Fig. 323

$$W = \frac{1}{2}\,\frac{CV^2}{n}.$$

The energy of the charge in the n similar jars is $1/n$ of the energy of one of the jars charged to the same available potential difference V between its two coatings.

518. Electric Strain.—The phenomenon of the residual charge may be best explained by considering the dielectric as the medium through which the induction takes place. The charging of a Leyden jar is accompanied by the straining of the glass. If the potential difference is raised to a sufficiently high value, the glass may be strained beyond its elastic limit and may give way with a disruptive discharge. The glass is shattered at the point through which the discharge takes place. In the case of air or other fluid dielectrics, such as insulating oils, the dielectric may be broken down by a disruptive discharge, but the damage is automatically repaired by the inflow of the insulating fluid.

By subjecting plate glass to powerful electrostatic stress and passing plane polarized light through it at right angles to the lines of force, Kerr discovered that the glass becomes double refracting, and is strained as if it were compressed along the lines of force. Quartz behaves in the same way. Kohlrausch and others have pointed out the analogy between the Kerr effect and the elastic fatigue of solids after they have been subjected to a twisting stress. A fiber of glass does not immediately regain its initial form when released from stress, but a slight set remains from which it slowly recovers. Its after recovery from distortion due to an electric charge sets free energy which is represented by the residual charge.

Hopkinson has shown that it is possible to superpose several residual charges of opposite signs. In the same way a glass fiber may be twisted first in one direction and then in the other, and the residual twists will appear in reverse order. No residual charge can be obtained from air condensers.

519. Electric Displacement. — Electricity exhibits some of the properties of an incompressible fluid. Electric charges show themselves only at the boundary between conductors and dielectrics. All cases of electrification are examples of the transfer of electricity. Hence Maxwell proposed his theory of *electric displacement.* It supposes that when an electromotive force acts on a dielectric, as in induction, electricity is displaced along the lines of induction. The electromotive force transfers electricity by distorting the dielectric. The strained dielectric by its elastic reaction produces a back electromotive force, and the discharge is a reverse electric transfer to restore the equilibrium.

520. The Dielectric Constant. — The density of the charges on the surfaces of the plates of a condenser, with a given potential difference between them, depends not only on their distance apart, but also on the facility with which the dielectric permits electric displacement.

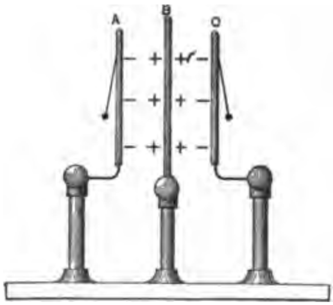


Fig. 324

Let A, B, C (Fig. 324), be three insulated conducting plates. From the back of A and C are suspended pith balls. Let B receive a positive charge and let A and C be charged negatively by induction. If they are touched with the finger, the pith balls collapse and remain in contact with the plates. If now A, for example, be moved nearer to B, both pith balls will diverge, the one on A with a positive charge and the one on C with a negative one.

Now replace A in the first position, with B charged as before and the pith balls not diverging. Interpose between A and B a thick plate of glass or sulphur. Both pith balls will again diverge as if A had been moved nearer to B. The effect is the same as the reduction of the thickness of air between the plates; the capacity of a condenser depends on the nature of the insulating medium between the plates.

The *dielectric constant* of a substance is the ratio of the capacity of a condenser with the substance as the dielectric to its capacity when the dielectric is air. To this constant of a dielectric Faraday gave the name *specific inductive capacity.*

The following are approximate values of K for some common dielectrics:

Glass	7.4	Ebonite	3.1
Mica	6.6	Petroleum	3.0
Sulphur	4.0	Paraffin	2.3
Shellac	3.2	Vaseline	2.0

521. Oscillatory Discharge. — It appears to have been first observed by Savary in 1824 that when a Leyden jar is discharged through a coil of insulated wire, consisting of a few turns and inclosing a sewing needle lying along its axis, the needle is not always magnetized in the expected direction. This curious fact was first explained by Joseph Henry in 1842. He discovered that the discharge of a Leyden jar consists of electric surges, first in one direction and then in the other, and that the energy of the discharge becomes smaller with each oscillation until it is all expended in heat. The last discharge strong enough to reverse the magnetism of the needle (which is confined to a superficial shell of the steel) determines its polarity.

In 1853 Lord Kelvin gave the mathematical theory of electric oscillations, and in 1858 Fedderson analyzed the spark of a small discharge into a string of images by a revolving mirror. These observations have since been con-

firmed by many observers. When the discharge is through a low resistance, the spark is a periodic phenomenon.

522. Electrical Resonance. — If two Leyden jars of the same capacity are attached to two similar discharge circuits, they should have the same period of oscillation and should therefore exhibit the phenomenon of resonance. Sir Oliver Lodge has found this to be true.

The two jars are connected to discharge circuits of the same size (Fig. 325); but while that of *A* is interrupted by a spark gap, that of *B* is closed and is adjustable by means of the slider *S*. If now the coatings of *A* are connected to the two electrodes of an influence machine, this jar discharges across the gap, and the oscillations at every discharge disturb the circuit of *B*, exciting in it feebler oscillations of the same period. If the two circuits are tuned to unison by moving the slider, the oscillations in *B* become sufficiently violent to make it overflow through the tin foil strip *c*, which comes over from the inner coating and nearly touches the outer one. The strip furnishes an easy overflow path, so that when the jars are near together and the two discharge circuits are parallel, every discharge of *A* is accompanied by a bright spark at the air gap *c*.

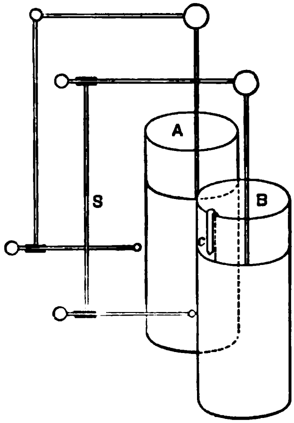


Fig. 325

Problems

1. Two Leyden jars are charged with quantities as 1 to 4. The tin foil surface of the second jar is twice as large as that of the first and the glass is half as thick. Find the relative energy of the two charges.

2. An insulated conducting sphere whose radius is 10 cm. is charged to a potential of 50 units. Find the number of units of charge.

3. If one of two conducting spheres, 25 cm. in diameter, be charged to a potential of 200 *c. g. s.* units (§ 500), and then be connected to the other sphere by means of a long thin wire, find the energy of the discharge between them.

4. A conducting sphere of 30 cm. radius is charged with 900 *c. g. s.* units. If it divides its charge with another insulated sphere of 10 cm. radius, what will be the charges on the two spheres?

5. Find the capacity of a spherical condenser, the radii of the opposed surfaces being 9½ and 10 cm., and the dielectric constant of the paraffin dielectric 2.3.

6. An insulated metal ball of 20 cm. radius and removed from all other conductors is charged with 200 *c. g. s.* units of electricity. What will be its potential if it is then surrounded by a smooth conducting shell of 20¼ cm. radius connected to earth?

7. Two concentric spheres of radii 10 cm. and 10½ cm. are separated by air and are charged to a potential difference of 800 units. Find the charge.

8. Find the capacity of a Leyden jar with coatings of 900 $cm.^2$, the glass 2 mm. thick, and the dielectric constant 7.

9. What is the energy of the charge if the jar of the last problem is charged to a potential of 1500 *c. g. s.* units?

10. What is the intensity of field between two parallel condenser plates which are ¼ cm. apart and differ in potential by 250 *c. g. s.* units?

V. ATMOSPHERIC ELECTRICITY

523. Lightning an Electrical Discharge. — Some of the early philosophers surmised that the lightning flash is an electrical discharge, but this view obtained little currency until after Franklin's suggestion to apply his discovery of the discharging power of points had been carried into effect. In 1752 D'Alibard, acting on Franklin's suggestion, erected an insulated iron rod 40 feet high and drew sparks from passing clouds. About the same time Franklin sent up his famous kite by means of a linen thread, during a passing storm, and held it by means of a silk ribbon between his hand and a key attached to the thread. After the thread had been wetted by the rain, sparks were drawn from the key and a Leyden

jar was charged. A year later Richmann of St. Petersburg was killed by lightning while experimenting with a rod similar to that of D'Alibard.

524. High Potential of Thunder Clouds. — The source of the electrification of the atmosphere and of clouds remains unsettled. But given ever so slight electrification of aqueous vapor, it is not difficult to account for the high potential of thunder clouds. Assume that each particle of water vapor has an initial charge, and that the larger globules are formed by the coalescence of smaller ones. If, for example, 1000 such particles unite to form a single large one, the diameter of the large drop will be only 10 times that of the component particles; but while the charge of the larger sphere is 1000 times that of the smaller ones, its capacity is increased only tenfold, since the capacity of a sphere is numerically equal to its radius. Its potential, however, has risen a hundred fold ($V = Q/C$). This rise of potential increases the inductive action between drops and the tendency to discharge from drop to drop in a cloudy atmosphere.

525. Effect of Electrification on Condensation. — A small ascending jet of water is resolved into drops, which describe different paths. By reason of the different velocities and directions of motion of the individual drops, they come into frequent collision and rebound. The influence of electrification on the recoil after collision is marked and interesting. The subject was investigated by Lord Rayleigh.

If the ascending jet is strongly electrified, the repulsion between the drops scatters them and prevents collision; but with feebler electrification the drops coalesce on impact and the stream is much smoother. The coalescence was shown to be due to slightly different degrees of electrification of the impinging particles. Their attraction and union appear to be due to induction, the resulting force of which is always an attraction.

The bearing of these results on the precipitation of aque-

ous vapor is obvious. Innumerable minute globules of water, feebly charged to different potentials, collide and coalesce into drops which descend by gravity. A slight degree of electrification of the atmosphere is therefore favorable to aqueous precipitation.

It is an observed fact of frequent occurrence that a vivid flash of lightning is quickly followed by a sudden downpour of rain. It is clearly impossible to tell which is antecedent to the other, the discharge or the condensation; for, while the flash reaches the observer first, light travels from the place of condensation in negligible time, and the discharge may therefore be subsequent to the sudden condensation. If the condensation occurs before the discharge, it is accompanied by a sudden rise of potential in the enlarged drops, leading to an electric discharge.

526. Lightning Flashes. — Lightning flashes are discharges between two oppositely charged conductors. They may occur between clouds, or between clouds and the earth. The rise of potential of a cloud causes an accumulation by induction in the earth underneath; and unless this accumulation is carried off by the silent action of points, when the electric stress in the air reaches a certain limiting value, the air as a dielectric breaks down with a sudden subsidence to equilibrium. Sir J. J. Thomson estimates the dielectric strength of the air under the usual conditions of pressure and temperature to be about 400 dynes per square centimeter. When the electric tension along lines of force becomes greater than this, a disruptive discharge takes place.

527. Discharge with Steady Strain. — If the stress in the dielectric is increased gradually, and the medium is finally strained to the point of rupture, a discharge takes place. Under these conditions a point will effect a silent discharge and afford protection. This condition Sir Oliver Lodge has named the "steady strain," and has illustrated it as follows:

A and *B* are the terminals of an influence machine, *L* is a Leyden jar, *T* and *T'* two metal plates connected with the two coatings of the jar as shown in Figure 326. On the lower plate are three conducting terminals of different heights.

When the machine is turned and the jar is charged, the stress increases gradually; but the pointed conductor, even when it is low compared with the others, prevents a discharge altogether. If the point be covered or removed, and the knobs be positive, long flashes may be obtained, but generally to the smaller ball unless it is much farther from the upper plate than the large one. The air breaks down at the weakest point, or where the stress is greatest, and this is at the surface of greatest curvature.

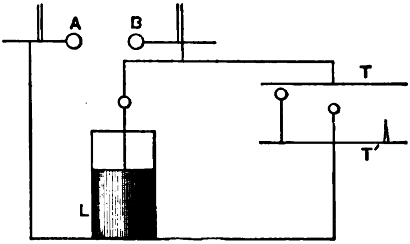


Fig. 326

528. Discharge with Impulsive Rush. — In the arrangement to illustrate the condition of "steady strain" the potential difference between the parallel plates increases gradually until the limit of the dielectric strength of the air is reached. Lodge has arranged a different experiment to illustrate the very sudden production of a high potential difference and a discharge with an "impulsive rush."

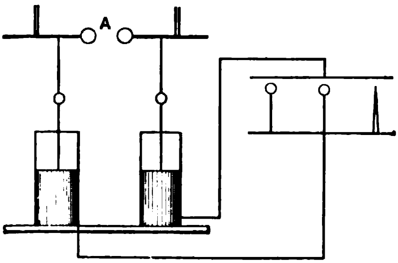


Fig. 327

The two Leyden jars in series (Fig. 327) stand on a wooden table or, better, have their outer coatings connected by a moist string. They charge slowly, the outer surfaces through the string, and finally discharge at *A*. The discharge between the inner coatings releases the charges on the outer ones, a violent rush takes place, producing a very sudden stress in the air between the plates, and one of the conductors is struck. The small ball

no longer protects the large one, nor does the point afford any certain protective influence. All three terminals are equally liable to be struck, if of the same height. In fact all three may be struck at once.

In this case the electric pressure is developed with such impulsive suddenness that the dielectric appears as liable to break down at one point as another. Such sudden rushes may occur when two charged clouds discharge into each other, and one overflows into the earth. The highest and best conducting objects may then be struck irrespective of points and terminals. The conditions of the discharge in the case of an impulsive rush are entirely different from those of the steady strain, and points are incompetent to afford protection by preventing the discharge.

529. Potential of the Air. — In clear weather the potential of the air is sometimes nearly as high as during a storm, but it shows smaller fluctuations. The value of the potential gradient found by McAdie at the Blue Hill Observatory was 0.013 electrostatic units per meter of elevation. On certain clear days the variation of potential with the elevation reached twice this value. During thunder storms the potential gradient may amount to 0.12 electrostatic units per meter.

By means of kites McAdie has shown that the potential difference in clear weather increases as the kite rises; and, further, that it is possible to obtain sparks from a perfectly cloudless sky, and generally at an elevation not exceeding 500 meters.

530. The Aurora. — The *aurora*, or polar light, is due to silent discharges in the upper regions of the atmosphere. Within the arctic circle it occurs almost nightly, and sometimes with indescribable splendor. Lemström has shown that the illumination of the aurora is due to currents of positive electricity passing from the higher regions of the atmosphere to the earth. In our latitude these silent streamers in the atmosphere are infrequent. When they do occur, they are accompanied by great disturbances of the earth's magnetism and by earth currents. Such magnetic disturbances occur at the same time in widely separated parts of the earth.

CHAPTER XVIII

ELECTRIC CURRENTS

I. VOLTAIC CELLS

531. An Electric Current. — When a condenser is discharged through a wire, there is produced in and around the wire a state called an *electric current*. Electrification is a condition of strain in the dielectric; an electric displacement through the discharging conductor rapidly relieves this strain. If the state of strain is reproduced by some agency as fast as it is relieved by the conductor, the result is a continuous electric current. The expression, "current of electricity," was introduced when electricity was regarded as a fluid flowing from higher to lower potential through a conductor, just as water flows through a pipe from a higher to a lower level.

To produce a continuous electric current through a conductor, a continuous potential difference must be maintained between its terminals. This may be accomplished by means of chemical energy, as in a voltaic cell; by the application of heat, as in a thermal couple; or through the agency of mechanical energy, as in the dynamo-electric machine. In all these cases the energy applied is converted, wholly or in part, into the energy associated with the transport of electricity under the electric pressure, which it is the function of the device or machine to establish and maintain. The energy of an electrostatic charge is potential energy; that of an electric current is kinetic energy.

The magnitude or strength of a current is the quantity of electricity passing any section of the conductor per second.

Quantity of electricity includes the sum of the positive transfer in one direction through the conductor and the negative in the other; for the flow of negative electricity in one direction has the same effect as that of an equal quantity of positive in the other. It is not necessary to distinguish between the two fluxes in opposite directions except through fluid conductors.

532. Simple Voltaic Cell. — The era of electric currents dates from Galvani's discovery in 1786 that violent muscular contractions are produced at the instant when a bimetallic arc of iron and copper connects the lumbar nerve and the crural muscle of a freshly killed frog. In the hands of his contemporary Volta this observation resulted in the discovery that a potential difference exists between two different metals, such as zinc and copper, when they are separated by moist cloth, damp muscle, or a conducting liquid. Hence the simple voltaic cell invented by Volta in the year 1800.

If a strip of zinc, amalgamated with mercury, be placed in sulphuric acid diluted with about twenty times its volume of water, bubbles of hydrogen will collect on the zinc, but the chemical action will soon apparently cease. No change is produced by placing a strip of clean copper in the same solution unless the two metals are connected directly or by means of some good conductor (Fig. 328). The acid then attacks the zinc, hydrogen is freely liberated *at the surface of the copper plate*, and a dense solution of zinc sulphate streams down from the zinc. The liquid product of the electrochemical reaction appears at the zinc plate and the gaseous product at the copper. As soon as the electrical connection between the two metals is interrupted, chemical action ceases and hydrogen is no longer disengaged.

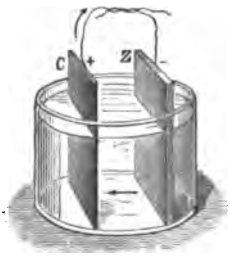


Fig. 328

Such a combination of two conductors, immersed in a liquid, called an *electrolyte*, which is capable of reacting chemically with one of them, is called a *voltaic cell.*

If an electroscope, terminating in a plate instead of a ball, be supplied with an extra loose plate insulated with shellac varnish, it becomes a condensing electroscope. If now the zinc strip of a voltaic cell be connected momentarily with the top plate or cover and the copper strip with the lower plate of the electroscope, when the cover is lifted with its insulating handle, the gold leaves will diverge slightly. Tested by means of an excited glass tube, they will be found to have a positive charge. If the zinc strip were connected with the lower plate, the gold leaves would diverge with a negative charge. Hence, the copper is *positive* and the zinc *negative*. The copper and zinc strips are called *electrodes*, the copper the *positive electrode* and the zinc the *negative electrode*.

When the two plates of a voltaic cell are joined by a conducting wire a number of new phenomena appear, which are characteristic of an electric current flowing through the conductor from the copper to the zinc, and through the liquid from the zinc to the copper.

533. The Circuit. — The circuit of a voltaic cell comprises the entire path traversed by the current, including the electrodes, and the liquid as well as the external conductor. *Closing the circuit* means completing it by joining the two electrodes by a conductor; *breaking* or *opening the circuit* is disconnecting them. When the circuit is closed, the zinc wastes away, and the energy of its union with the acid is in part given out by degrees as the energy of the electric current, which may be made to do work or to generate heat.

534. Electrochemical Action in a Voltaic Cell. — The theory of dissociation furnishes an explanation of the manner in which an electric current is conducted through a liquid. It is briefly as follows: When a salt or an acid, such as hydro-

chloric acid (HCl), is dissolved in water, some of the molecules split into two parts ($\overset{+}{H}$ and $\overset{-}{Cl}$, for example), one part having a positive electrical charge and the other a negative one. The two parts of the dissociated substance with their electrical charges are called ***ions.*** (The term is from a Greek word meaning *to go;* ions are travelers or carriers.) An electrolyte is a compound capable of such dissociation into ions. It conducts electricity only by means of the migration of the ions resulting from the splitting in two of the molecules. The separate ions convey their charges with a slow and measurable velocity through the liquid. Electropositive ions, such as zinc and hydrogen, carry positive charges in one direction; electronegative ions, such as chlorine and "sulphion" (SO_4), carry negative charges in the opposite direction; the sum of the two kinds of charges carried through the liquid per second is the measure of the current.

The active components in a voltaic cell, set up with hydrochloric acid, may be represented as follows:

$$\begin{array}{ccccccc} & \overset{+}{H} & \overset{+}{H} & \overset{+}{H} & \overset{+}{H} & \overset{+}{H} & \\ Zn & & & & & & Cu \\ & \overset{-}{Cl} & \overset{-}{Cl} & \overset{-}{Cl} & \overset{-}{Cl} & \overset{-}{Cl} & \end{array}$$

Immediately after the circuit has been closed this becomes

$$\begin{array}{ccccccc} & \overset{++}{Zn} & & \overset{\longrightarrow}{\overset{+}{H}} & \overset{+}{H} & \overset{+}{H} & \\ Zn & & & & & & 2H,\ Cu \\ & \overset{-}{Cl} & \overset{-}{Cl} & \underset{\longleftarrow}{\overset{-}{Cl}} & \overset{-}{Cl} & \overset{-}{Cl} & \end{array}$$

Zinc goes into solution as zinc chloride ($ZnCl_2$), and hydrogen appears as free hydrogen gas at the copper plate. Zinc ions crowd out hydrogen ions, while the positive and negative charges brought to the copper and zinc plate respectively reunite as a current through the external conductor.

535. Electromotive Force.— A voltaic cell is an electric generator. It is analogous to a rotary pump which produces a difference of pressure between its inlet and its outlet. Such a pump may cause water to circulate through a system of horizontal pipes against friction. In any portion of the system the force producing the flow is the difference of water pressure between the two ends of this portion. The force is all applied at the pump, and it produces a pressure throughout the whole circuit.

A voltaic cell generates electric pressure called *electromotive force.* It produces electric pressure to set electricity flowing. The seat of the electromotive force (E. M. F.) in a voltaic cell is at the contact of the dissimilar substances in the cell, and chiefly at the contact surfaces between the electrodes and the electrolyte. The E. M. F. of any form of voltaic cell depends on the materials employed, and it is entirely independent of the shape and size of the electrodes; it is modified by oxidation and by the concentration of the solutions.

The E. M. F. of a cell is the measure of the work required to transport unit quantity of electricity around its entire circuit. Work is required to effect this transfer, because all conductors offer resistance to the passage of a current, and back or opposing E. M. F.'s are sometimes present. The energy expended against resistance goes to heat the conductor.

536. Difference of Potential. — The difference of potential between two points on the external conducting circuit is the work done in carrying unit quantity of electricity from one point to the other. If E denotes this potential difference and Q the quantity conveyed, then the work done is the product EQ. But the quantity conveyed by a conductor per second is the strength of current I. The energy transformed, therefore, when a current I flows through a conductor, under electric pressure or potential difference of E units between its ends, is EI ergs per second.

537. Polarization.—When the circuit of a simple voltaic cell is closed, the current falls off rapidly in intensity, and at length almost ceases to flow. The hydrogen covering the copper plate as a film produces a state known as the *polarization* of the cell. The accumulation of ions on the electrodes changes the potential difference between them and the electrolyte. The hydrogen on the positive electrode not only introduces more resistance to the flow of the current, but it diminishes the electromotive force to which that flow is due, by setting up an opposing difference of potential.

538. Remedies for Polarization. — Any device that will prevent the liberation of hydrogen and its deposit on the positive electrode is a remedy for polarization. These remedies are mostly chemical, as illustrated by the following experiment, which was devised by Mr. D. H. Fitch about 1879:

Set up a cell by placing in a small glass jar enough clean mercury to cover the bottom, and filling it with a saturated solution of common salt. Hang a plate of zinc in the liquid, and thrust into the mercury the exposed end of a platinum wire, sealed into a glass tube, to connect with the mercury as the positive electrode. Close the circuit through some simple current indicator, such as a common telegraph sounder of low resistance. The armature will be drawn down strongly at first; but in the course of a minute or two the magnet will release it, showing that the current is very greatly weakened by polarization. The sodium released on the surface of the mercury attacks the water, producing sodium hydroxide and hydrogen.

Keeping the circuit closed, drop into the cell a very small piece of mercuric chloride ($HgCl_2$) no larger than the head of a pin. The armature of the sounder will be drawn down suddenly, showing recovery of the cell from polarization. The mercuric chloride furnishes chlorine ions which unite with the hydrogen ions on the surface of the mercury electrode and reduce the polarization. The chlorine will be exhausted in a few minutes, and polarization will again ensue.

539. The Daniell Cell. — The Daniell cell illustrates the chemical method of avoiding polarization by replacing the hydrogen ions by others, such as copper or mercury, which do not produce polarization when they are deposited on the positive electrode. In the Daniell cell the copper plate is

surrounded with a saturated solution of copper sulphate ($CuSO_4$), so that copper instead of hydrogen is deposited on the copper electrode.

A zinc bar Z (Fig. 329) is immersed in acidulated water, or in a dilute solution of zinc sulphate ($ZnSO_4$), in an unglazed earthenware cup; the copper plate C is a cylinder of sheet copper surrounded with a saturated solution of copper sulphate. Spare crystals of this salt are added to keep the solution saturated during the action of the cell. The porous cup allows the ions to pass through its pores, but prevents a rapid admixture of the two electrolytes.

Fig. 329

Both electrolytes undergo partial dissociation into ions; and when the circuit is closed the electropositive zinc and copper ions both travel toward the copper electrode. The zinc ions do not reach it, because zinc in copper sulphate solution replaces the copper, forming zinc sulphate. The relation of the several ions in the Daniell cell may be represented graphically as follows:

$$\begin{array}{ccccccc} & & & \longrightarrow & & & \\ & \overset{++}{Zn} & \overset{++}{Zn} & \overset{++}{Cu} & \overset{++}{Cu} & \overset{++}{Cu} & \\ Zn & & & & & & Cu \\ & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & \\ & & & \longleftarrow & & & \end{array}$$

As soon as the circuit is closed, $ZnSO_4$ is formed at the zinc electrode and copper is deposited on the copper electrode.

The so-called *gravity cell* is a Daniell cell, and in it advantage is taken of the difference in density of the two sulphate solutions to keep them in a measure separate. The copper electrode is placed in the bottom of the

jar with the $CuSO_4$ solution, and the lighter dilute $ZnSO_4$ solution floats on top with the zinc suspended in it (Fig. 330). The cell must be kept at work to prevent the diffusion of the $CuSO_4$ upward as far as the zinc.

Fig. 330

540. The Leclanché Cell. — The Leclanché cell belongs to a class containing a depolarizer, the office of which is to supply oxygen to unîte with the hydrogen and form water. In this cell the depolarizer is solid manganese dioxide (MnO_2). It is a zinc-carbon couple with a saturated solution of ammonium chloride (NH_4Cl) as the electrolyte. The carbon electrode is packed in a porous cup with the manganese dioxide in granules mixed with broken carbon to increase the conductance. The zinc is a rolled rod about one centimeter in diameter. Figure 331 shows a complete cell. The porous cup in this particular form has a flange resting on the top of the glass jar. This closes the jar and prevents evaporation.

Fig 331

If the circuit be kept closed for several minutes, the accumulation of hydrogen on the carbon plate produces some polarization; but when the circuit is opened again, the depolarizer slowly removes the hydrogen and the cell recovers its normal E.M.F. No serious local chemical action takes place on open circuit; the cell will stand without material waste for months. It is this characteristic that makes it suitable for many domestic purposes, or for intermittent service.

The initial relation of the ions in the Leclanché cell may be represented as follows:

$$
\begin{array}{cccccc}
 & & \longrightarrow & & & \\
 & \overset{+}{NH_4} \quad \overset{+}{NH_4} & & \overset{+}{NH_4} \quad \overset{+}{NH_4} & & \\
Zn & & & & & C \\
 & \overset{-}{Cl} \quad \overset{-}{Cl} & & \overset{-}{Cl} \quad \overset{-}{Cl} & & \\
 & & \longleftarrow & & &
\end{array}
$$

After the circuit is closed the first step is

$$
\begin{array}{cccccc}
 & & \longrightarrow & & & \\
 & \overset{++}{Zn} & & \overset{+}{NH_4} \quad \overset{+}{NH_4} & & \\
Zn & & & & & 2NH_3,\ 2H,\ C \\
 & \overset{-}{Cl} \quad \overset{-}{Cl} & & \overset{-}{Cl} \quad \overset{-}{Cl} & & \\
 & & \longleftarrow & & &
\end{array}
$$

One atom of zinc unites with two atoms of chlorine, or zinc has a valence of two and carries two ionic charges as compared with hydrogen or chlorine. Ammonia is released and the free hydrogen unites with oxygen from the manganese dioxide in accordance with the chemical reaction

$$2\,H + 2\,Mn\,O_2 = H_2O + Mn_2\,O_3.$$

The so-called "dry cells" in common use are substantially Leclanché cells, and the equations above describe the electrochemical action going on in them.

541. Standard Cells.—For many methods of precise electrical measurement and for purposes of standardizing instruments, a definitely known electromotive force is necessary. This requirement is met by a *standard cell*, which is never employed to furnish a current, but only a known potential difference to be very precisely balanced against some other potential difference to be measured.

The standard cell chiefly in use for many years was invented by Latimer Clark and known as the Clark cell. The negative electrode is a 10 per cent amalgam of zinc in

a neutral saturated solution of zinc sulphate; the positive electrode is pure mercury covered with a paste of mercurous sulphate and zinc sulphate. The cell (Fig. 332) is nearly filled with zinc sulphate crystals and the saturated solution of this salt. Both legs of the cell are hermetically sealed.

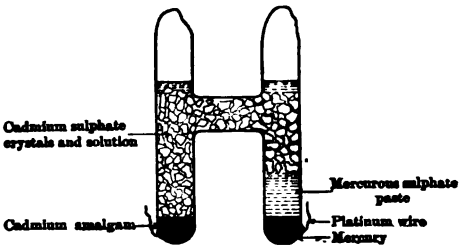


Fig. 332

The Clark cell has now practically been replaced by the "Weston Normal Cell" because of its great advantage over the Clark in that its change of E. M. F. with temperature is only one thirtieth as great. It is set up in the same manner as the Clark, except that cadmium and cadmium sulphate take the place of zinc and zinc sulphate. The negative electrode is an amalgam containing from 12 to 12.5 per cent of cadmium.

The E. M. F. of the Weston cell is given by the equation

$$E = 1.0184 - 0.0000406\,(t - 20°) - 0.00000095\,(t - 20°)^2.$$

(The E. M. F., E, is in volts equal to 10^8 *c. g. s.* electromagnetic units as explained in § 584.)

542. Effects of Heat on Voltaic Cells. — Two different effects are produced by heating a voltaic cell. The resistance of the liquid to the passage of the current is reduced, and the E. M. F. suffers a small change, either an increase or a decrease.

Professor Daniell found that a larger current flowed from his cell when he heated it to 100°. This increase was due to the fact that the relative decrease in the internal resistance of the cell was larger than the relative decrease in its E. M. F. The curve in Figure 333 shows the relation between

the internal resistance and the temperature of a Daniell cell between 13° and 68°. The resistance at the upper temperature is reduced to less than one half its initial value.

The temperature coefficient of a Daniell cell is only about 0.007 per cent; that is, the E. M. F. falls 0.007 volt for a rise of temperature of 100 degrees.

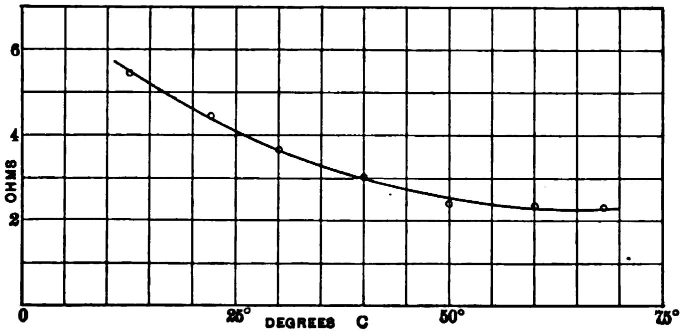


Fig. 333

II. ELECTROLYSIS

543. Electrolytes.—When an electric current passes through metals or carbon (often called conductors of the first class), heat is generated, but the conductor undergoes no change in chemical composition or physical state, except the change caused by the rise in temperature. Many liquids also conduct electricity, notably a large number of chemical compounds, either fused or in solution (often called conductors of the second class); the passage of electricity through these is always accompanied by chemical decomposition. For this reason the process is called *electrolytic conduction* or *electrolysis*, and the substance decomposed, an *electrolyte*.

The conductors of the first class by which the current enters and leaves the electrolyte are called *electrodes*, as in the case of voltaic cells. The current enters the electrolyte by the *anode* and leaves it by the *cathode*. The electrolyte is

split in two, the two parts migrating in opposite directions; the part migrating toward the anode is called the *anion*, and the other part, migrating toward the cathode, is the *cation*. All these terms are due to Faraday.

The initial products of the electrolysis are not always set free at the electrodes, because secondary chemical reactions sometimes take place between them and the electrodes or the solvent. Such is the case in a voltaic cell when the hydrogen is oxidized by the depolarizer. These secondary reactions are sometimes termed "secondary electrolysis."

544. Electrolysis of Copper Sulphate. — Copper sulphate presents one of the simplest examples of electrolysis. If the electrodes are copper, the passage of an electric current simply transfers copper from the anode to the cathode. The copper deposited on the cathode is very pure, and the process is now employed on a colossal scale for refining copper.

When copper sulphate is dissolved in water, it undergoes dissociation to some extent. If electric pressure is applied to the solution through the electrodes, the electropositive ions ($\overset{++}{Cu}$) are set in motion from higher to lower potential, and the electronegative ions ($\overset{--}{SO_4}$) in the opposite direction. The Cu ions are driven against the cathode, and, giving up their charges, become metallic copper. The SO_4 ions go to the anode, and, giving up their charges, unite with the copper of the anode, forming copper sulphate. Copper is thus removed from the anode as fast as it is deposited on the cathode.

The relation of the ions to the current may be represented graphically as follows:

$$
\begin{array}{cccccccc}
 & & & & \longrightarrow & & & \\
 & & \overset{++}{Cu} & \overset{++}{Cu} & \overset{++}{Cu} & \overset{++}{Cu} & & \\
\text{Anode} & Cu & & & & & Cu & \text{Cathode} \\
 & & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & & \\
 & & & & \longleftarrow & & &
\end{array}
$$

The passage of a current through an electrolyte is accomplished in the same manner, whether it is in a voltaic cell or in an electrolytic cell. Molecules not dissociated are electrically neutral and take no part in the transfer of electricity.

Since metallic copper is deposited from the solution on the cathode, and SO_4 migrates from the cathode toward the anode, the concentration of the solution at the cathode is diminished by electrolysis. At the anode, on the other hand, the concentration is increased. The change in concentration at the cathode may be shown by arranging the anode at the bottom of a large glass tube about 20 cm. long, filling with saturated $CuSO_4$ solution, and electrolyzing for some time with large current density. The liquid near the cathode will become nearly or quite colorless.

545. Electrolysis of Water. — Water appears to have been the first substance decomposed by an electric current; but it was a mooted question for about three quarters of a century whether the decomposition of the water was a result of a primary electrolysis or only that of a secondary chemical reaction. It is now known that pure water does not conduct an appreciable current of electricity; but if it is acidulated with a small quantity of sulphuric acid, it is decomposed as a secondary action.

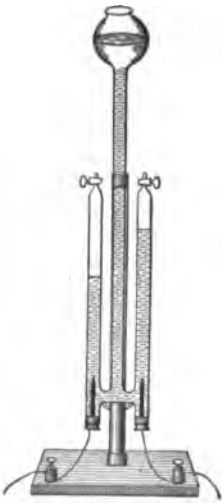


Fig. 334

In Hofmann's apparatus (Fig. 334) the acidulated water is poured into the bulb at the top, and the air escapes by the glass taps until the tubes are filled. The taps are then closed, and if connection is made with a battery of three or more cells in series, bubbles of gas will be liberated on the pieces of platinum foil at the bottom. The gases collecting in the tubes may be examined by allow-

ing them to escape through the taps. Oxygen will be found at the electrode by which the current enters the apparatus, and hydrogen at the other; that is, oxygen collects at the anode and hydrogen at the cathode.

The electrochemical action may be represented as follows:

$$\text{Anode}\quad Pt\quad \begin{array}{c}\longrightarrow \\ 2\overset{++}{H}\ \ 2\overset{++}{H}\ \ 2\overset{++}{H}\ \ 2\overset{++}{H} \\ \\ \overset{--}{SO_4}\ \ \overset{--}{SO_4}\ \ \overset{--}{SO_4}\ \ \overset{--}{SO_4} \\ \longleftarrow \end{array}\quad Pt\quad \text{Cathode}$$

Hydrogen is set free at the cathode, while at the anode the secondary electrolysis is

$$H_2O + SO_4 = H_2SO_4 + O.$$

The primary electrolysis is that of the dissociated sulphuric acid; the water is decomposed at the anode by the SO_4. As often as one atom of oxygen is set free at the anode, two of hydrogen are liberated at the cathode. The volume of the hydrogen is not exactly twice that of the oxygen, because the latter is more soluble in water than the former, and about one per cent of it is evolved in the denser form of ozone; on the other hand, more hydrogen than oxygen is absorbed or occluded by the platinum electrodes.

546. Electrolysis of Sodium Sulphate. — When a salt of one of the alkali metals is electrolyzed between platinum electrodes, secondary reactions take place at both electrodes. If the electrolyte is sodium sulphate, free sodium cannot exist at the cathode, but it unites with water, forming sodium hydroxide; at the anode the SO_4 decomposes water and liberates oxygen.

Fill a V-shaped tube (Fig. 335) two thirds full with a solution of sodium sulphate, colored with the extract of purple cabbage or purple violets. Close the ends with corks and thrust through them platinum

wires terminating within the tube in strips of platinum foil. When the current is passed, the liquid turns red at the anode and green at the cathode, showing the presence of an acid at the former and an alkali at the latter. Stop the flow of current and mix the liquids at the two electrodes; both the red and green color will disappear with the restoration of the faint purple, demonstrating that the acid and the alkali are produced in chemically equivalent quantities.

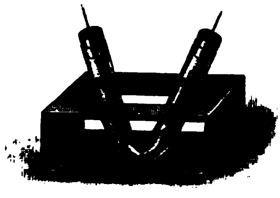


Fig. 335

The final result is the liberation of oxygen at the anode and hydrogen at the cathode.

$$\begin{array}{cccccccc} & & & \longrightarrow & & & & \\ & & 2\overset{++}{Na} & 2\overset{++}{Na} & 2\overset{++}{Na} & 2\overset{++}{Na} & & \\ \text{Anode} & Pt & & & & & Pt & \text{Cathode} \\ & & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & \overset{--}{SO_4} & & \\ & & & \longleftarrow & & & & \end{array}$$

$$H_2O + SO_4 = H_2SO_4 + O. \qquad 2\,Na + 2\,H_2O = 2\,NaOH + 2\,H.$$

If mercury is used as the cathode in an aqueous solution of sodium sulphate or sodium chloride, the separated metallic sodium amalgamates with the mercury, and only a little sodium hydroxide is formed. When fused sodium chloride (common salt) is electrolyzed between a carbon anode and molten lead as the cathode, the metallic sodium alloys with the lead. These two facts form the basis of two methods of manufacturing caustic soda by electrolysis.

547. Electrolysis of Lead Acetate. — Place the solution, which may be made clear by the addition of a little acetic acid, in a flat glass tank and electrolyze between two lead wires as electrodes. The lead separated from the clear solution will be deposited on the cathode in the form of shining crystals, which grow rapidly, giving rise to the "lead tree." If the process is not conducted too rapidly, these crystals assume very beautiful forms. The lead goes into solution at one electrode and comes out of solution at the other.

After a few minutes reverse the current; the first crystalline deposit will gradually disappear, and another one will

form on the other lead wire. In this way the disappearance of the lead at the anode and its appearance out of a clear solution at the cathode may be observed at the same time. The reaction is exactly the same as that of copper sulphate between copper electrodes.

548. Faraday's Laws of Electrolysis. — When several electrolytic cells are joined in series, in the manner shown in Figure 336, the same current passes through all of them. When the several cathodes are weighed, it is found that the masses of the different elements deposited are directly as their atomic weights and inversely as their valences. Thus, silver has a valence of one and an atomic weight of 107.94, while copper has a valence of two and an atomic weight of 63.4. The masses of the two deposited by the same quantity of electricity are therefore as $107.94 : \frac{63.4}{2}$. The ratio of the atomic weight of an element to its valence is known as its "chemical equivalent."

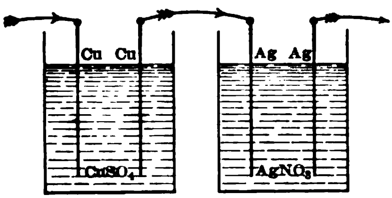


Fig. 336

From an extensive study of electrolytic phenomena Faraday deduced the following laws:

I. *The mass of an electrolyte decomposed by an electric current is directly proportional to the quantity of electricity conveyed through it.*

II. *When the same quantity of electricity is conveyed through different electrolytes, the masses of the different ions set free at the electrodes are proportional to their chemical equivalents.*

Thus, if the same current be passed through a series of electrolytic cells, in which it liberates hydrogen, chlorine, copper, and silver, for every gram of hydrogen set free,

35.46 gm. of chlorine, 31.7 of copper, and 107.94 of silver will be separated.

The *electrochemical equivalent* of an element is the number of grams of it deposited by the passage of unit quantity of electricity. When a current has unit strength, unit quantity flows through any cross section of the conductor per second. Faraday's laws may then be combined in the one statement that the number of grams deposited by the passage of a constant current through an electrolyte is equal to the continued product of the strength of the current (in amperes), the time in seconds during which it flows, and the electrochemical equivalent of the element. If z denotes the electrochemical equivalent and I the current strength, then in symbols

$$M = Itz. \tag{93}$$

The *coulomb* is the unit of quantity in the electromagnetic system of measurement to be described later. The current strength is an *ampere* when the quantity transmitted is at the rate of a coulomb per second. The electrochemical equivalent of silver has been determined by absolute measurement of the current in amperes and observing the time. The accepted value is 0.001118 grams per coulomb.

549. Charge conveyed by the Gram Equivalent. — The *gram equivalent* of an element is the number of grams equal to its chemical equivalent. Thus, the gram equivalent of silver is 107.94; of copper, 31.8; of zinc, 32.7. Since one coulomb of electricity separates 0.001118 gm. of silver, it will require 107.94/0.001118 = 96,540 coulombs to separate one gram equivalent. An equivalent statement of Faraday's second law is that the gram equivalent of any element transports the same quantity of electricity or carries the same charge; this quantity is 96,540 coulombs.

Since the gram equivalents of univalent ions are proportional to their atomic weights, it follows that the charge carried by each ion is the same for all univalent atoms. If e is this charge in coulombs and m the mass

in grams of an atom of hydrogen (the gram equivalent of which is one), then

$$e = 96{,}540\ m,$$

if the atom of hydrogen is the mass of its ion. The charge carried by a divalent ion is $\pm 2e$; by a trivalent ion, $\pm 3e$, etc.

The *c. g. s.* electromagnetic unit of quantity is ten times the coulomb. If, then, e is in *c. g. s.* units,

$$e = 9{,}654\ m.$$

The ratio e/m in the case of hydrogen in liquid electrolysis is thus approximately 10^4.

550. Polarization of an Electrolytic Cell. — If the two platinum electrodes of Hofmann's apparatus be connected to a sensitive galvanometer (§ 588) immediately after they have been used for the electrolysis of sulphuric acid, it will be found that some energy has been stored, for the cell will furnish a current. The chemical and electrical functions are now reversed; the hydrogen and oxygen in contact with the electrodes unite to form water, and a reverse current flows through the cell. The apparatus may be set up as in Figure 337. *B* is the battery for the electrolysis of the sulphuric acid. Hydrogen accumulates in the tube *H* and oxygen in the tube *O*. Let the two-point switch *S* be now turned so as to cut off the battery and to join the electrolytic cell to the galvanometer *G*. The needle will be sharply deflected by the current from the Hofmann's apparatus. To determine its direction, a thermal couple (§ 571), consisting of a copper and an iron wire soldered together and placed in the circuit of the galvanometer at *T*, is convenient. When such a couple is slightly heated, a current passes across from *Cu* to *Fe*. It may be tried before charging the electrolytic cell, and the direction of the deflection of the galvanometer should be noted. It

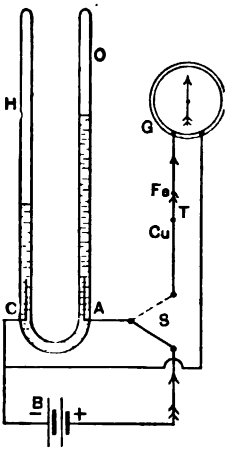


Fig. 337

will be found that the current due to the polarization of the electrolytic cell flows out from *A* and in at *C*, or in the reverse direction to the current which separates the gases oxygen and hydrogen. The E. M. F. of polarization is therefore a back or resisting E. M. F.

551. The Storage Cell. — If the platinum electrodes of the sulphuric acid electrolytic cell be replaced by lead, we have the Planté storage cell, which is the basis of all modern storage batteries.

Attach two lead plates, to which are soldered copper wires, to opposite sides of a block of dry wood, and immerse them in a twenty per cent solution of sulphuric acid (Fig. 338). Pass a current through the cell for a few minutes. The oxygen liberated at the anode will oxidize the lead, forming a dark brown or chocolate-colored coating of peroxide of lead. An ordinary electric house bell may be connected to the cell by a switch, as in Figure 337. When the switch is turned, cutting off the charging battery and connecting the lead electrolytic cell with the bell, the latter will ring vigorously for a few seconds. The operation may be repeated, showing that energy is stored in the cell by the process of electrolysis. Planté subjected his cells to repeated charging in opposite directions, so that both plates should be modified to an appreciable depth by alternate oxidation and reduction.

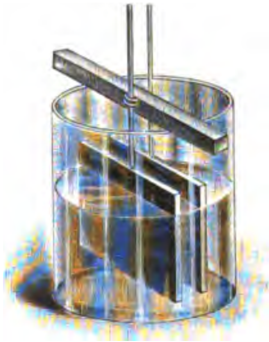


Fig. 338

In most modern storage cells the lead plates, cast or rolled in the form of grids, are filled with lead oxides called the active material. These oxides are changed into peroxide at the anode, and reduced by hydrogen to spongy lead at the cathode, during the process of charging. The chemical reactions of the storage cell are complex and to some extent undetermined. Sulphuric acid is formed during the charging of the cell and disappears during the discharge. Some sulphate of lead is also formed during the discharge, and may be

reduced by hydrogen with slow charging. The electrode which is the anode when charging and the cathode when discharging, is called the positive plate. Figure 339 shows a complete storage cell containing one positive plate between two negatives.

Fig. 339

Many voltaic cells are reversible; that is, when a reverse current is sent through them, the electrochemical reactions in them are all reversed as compared with those taking place when the current is in the normal direction, and energy is stored by means of electrolysis. A storage cell is a reversible one specially designed to store the applied electrical energy, so that it can be recovered at a subsequent time in the form of a current.

Problems

1. How much silver will a current of one ampere deposit in an hour?

2. The silver deposited on the cathode in the electrolysis of silver nitrate was 2.8095 gm. in 45 minutes. What was the average current?

3. How much copper will a current of 1000 amperes deposit per hour?

4. What current will deposit 0.5 gm. of copper in an hour?

5. If the current in the last problem is furnished by a Daniell battery, how much zinc will go into solution and how much copper will be deposited in each cell during the hour?

6. If doubling the current through a given circuit requires doubling the number of similar cells in series, how much more zinc will be consumed per second in the entire enlarged battery?

III. OHM'S LAW AND ITS APPLICATIONS

552. Ohm's Law. — When an electromotive force produces a current flow in any given circuit, the strength of the current is always proportional to the value of the electromotive force. This law was discovered by Dr. Ohm of Berlin in 1827, and it has since been known as Ohm's Law.

If E be the potential difference between two points of a conductor and I the numerical value of the current flowing through it, then if suitable units be chosen,

$$E = RI, \tag{94}$$

where R is a proportionality factor called the *resistance* of the conductor. The resistance is independent of the value and direction of the current flowing, and depends only on the material of the conductor, its length and sectional area, and its temperature.

Equation (94) is an expression of Ohm's law. It is usually written in the equivalent form

$$I = \frac{E}{R}. \tag{95}$$

If the practical units now adopted by international agreement be employed, Ohm's law may be expressed without symbols as follows: The number of *amperes* flowing between two points of a circuit is equal to the number of *volts* of potential difference divided by the number of *ohms* of resistance between the same points.

When Ohm's law is applied to the entire circuit, which may contain several sources of E. M. F. of different signs, and both metallic and electrolytic resistances, attention must be paid to the signs of the electromotive forces. If, for example, there are several voltaic cells in the circuit, some of them may be connected in the wrong direction so that they oppose the flow of current; or the circuit may include electrolytic or storage cells or motors, which offer a resistance to the flow of the current in the form of a counter E. M. F. All such electromotive forces must be regarded as negative.

There may be also a number of consecutive resistances in series, but resistance is not a directed quantity, for it restricts the flow of the current, whether it be in one direction or the other.

Ohm's law may then be written

$$I = \frac{E_1 + E_2 + E_3 + \cdots}{R_1 + R_2 + R_3 + \cdots} = \frac{\Sigma E}{\Sigma R}, \tag{96}$$

where each E. M. F. must be taken with its proper sign.

553. Resistance. — Resistance has already been defined as a proportionality factor in the equation expressing Ohm's law. It is moreover the property of a conductor by virtue of which the energy of a current is converted into heat. It is independent of the direction of the current, and the conversion of electrical energy into heat occasioned by it is an irreversible one; that is, there is no tendency for the heat energy to revert to the energy of an electric current.

The practical unit of resistance is the *ohm*. It is represented by *the resistance offered to an unvarying current by a thread of mercury at the temperature of melting ice*, 14.4521 *gm. in mass, of uniform cross-sectional area, and of a length of* 106.3 *cm.* This definition is equivalent to saying that the cross section of the thread of mercury is one square millimeter, but it avoids any assumption respecting the density of mercury.

554. Resistivity. — The resistances of diverse conductors are found to conform to the following laws:

I. *The resistance of a uniform conductor is directly proportional to its length.*

II. *The resistance of a uniform conductor is inversely proportional to its cross-sectional area.*

III. *The resistance of a uniform conductor of given length and cross section depends upon the material of which it is made.*

This property of a conductor, which determines its resistance and depends on the nature of its material, is called *resistivity*.

In symbols the three laws of resistance may be expressed by the equation

$$R = \frac{\rho l}{a}, \qquad (97)$$

in which ρ is the measure of the resistivity. Obviously $\rho = R$ when both l and a are unity; or, if l is in centimeters and a in square centimeters, ρ is numerically equal to the resistance of a conductor 1 cm. long and 1 cm.^{2} cross-sectional area. Resistivity is expressed in *c.g.s.* units; the ohm is equal to 10^9 *c.g.s.* units of resistance.

Table

Resistivities in *c.g.s.* Units at 0°

Lead	20,380	Iron	9,065
Thallium	17,633	Zinc	5,751
Tin	13,048	Magnesium	4,355
Nickel	12,323	Aluminum 99 %	2,568
Platinum	10,917	Gold	2,197
Palladium	10,219	Copper	1,561
Cadmium	10,023	Silver	1,468

555. Conductance. — The inverse of a resistance is a *conductance.* A conductor whose resistance is r ohms has a conductance equal to $1/r$. When a number of conductors are joined in parallel, the conductance of the whole is the sum of

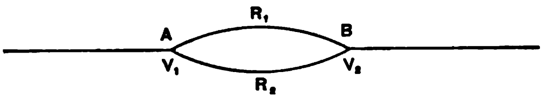


Fig. 340

the conductances of the several branches. If two conductors of resistances R_1 and R_2 are connected in parallel between the points A and B (Fig. 340), and if V_1 and V_2 are the potentials of A and B respectively, putting $V_1 - V_2 = E$, we have by Ohm's law,

$$\frac{E}{R} = \frac{E}{R_1} + \frac{E}{R_2}.$$

The first member of this equation is the whole current flowing, and this is equal to the sum of the currents through the two branches. R is the combined resistance of the two conductors in parallel. Hence

$$\frac{1}{R}=\frac{1}{R_1}+\frac{1}{R_2}.$$

From the last equation,

$$R=\frac{R_1R_2}{R_1+R_2} \tag{98}$$

556. Relation between Resistivity and Temperature. — The resistivity of all pure metals and of most alloys increases with rise of temperature. The resistivity of carbon and of electrolytic conductors decreases when the temperature rises. Thus, the resistance of a carbon incandescent lamp filament is only about half as great at normal incandescence as when cold. Solutions of zinc sulphate and of copper sulphate have a temperature coefficient somewhat over 0.02, or 2 per cent per degree C.

The curves of Figure 341 show the variation of resistivity with temperature for a number of pure metals. All these curves tend toward a point of convergence of zero resistivity at a temperature near the absolute zero of $-273°$ on the centigrade scale. Over comparatively short ranges of temperature these curves are approximately straight lines, or the change in resistivity is nearly proportional to the change in temperature. It follows that the resistivity at any temperature $t°$ may be expressed by the relation

$$\rho_t=\rho_o\,(1+at), \tag{99}$$

in which ρ_t is the resistivity at any temperature $t°$, ρ_o that at 0°, and a is the *temperature coefficient.* The same expression holds for the resistance of any particular conductor. It will be seen from the diagram that the temperature coefficient of platinum is very nearly a constant over a wide range of

temperature. It is the same as the coefficient of expansion of a perfect gas, 0.00367 (§ 380). In fact, the relation between the resistivity of platinum and its temperature might be used to define the absolute zero of temperature just as well as the expansion of a perfect gas. The temperature coefficient of the other pure metals is a quantity of the same order of magnitude as that of platinum, but in general a trifle larger.

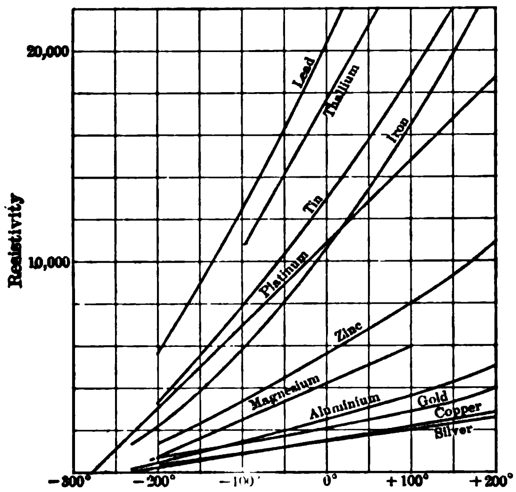


Fig. 341

The temperature coefficient of alloys is smaller than that of pure metals. German silver has a coefficient only about one tenth as great as that of copper. Manganin, an alloy of copper, manganese, and nickel, has a temperature coefficient nearly equal to zero; at certain temperatures manganin has a small negative coefficient; that is, its resistivity diminishes slightly as the temperature rises.

557. Loss of Potential Proportional to Resistance. — If V_1 and V_2 are the potentials of two points on a conductor carrying a current I, by Ohm's law (94),

$$V_1 - V_2 = RI.$$

It is obvious from this relation that the potential difference between any two points of a conductor through which a constant current is flowing is proportional to the resistance between them, provided the conductor is not the seat of an E. M. F. Even when electromotive forces are encountered, the *loss* of potential, when a given current flows through a conductor, is proportional to the resistance of the conductor.

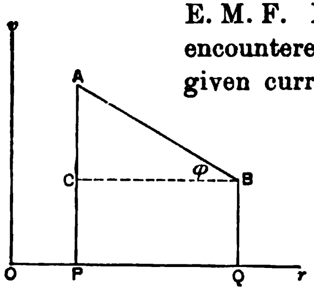


Fig. 342

Let distances measured along Or denote resistances (Fig. 342), and those along Ov potential differences. Lay off AP equal to V_1 and BQ to V_2; also the distance PQ stands for the resistance R between the points A and B on the conductor. Then AC equals $V_1 - V_2$, and the slope of the line AB represents the rate at which the potential drops per unit length along the resistance R. Moreover, since

$$\tan\phi = \frac{AC}{BC} = \frac{E}{R} = I,$$

it is evident that the tangent of the angle of slope equals the strength of the current.

The principle that the loss of potential is equal to the resistance passed over, *when the current is constant*, is one of very frequent application in electrical measurements.

558. Wheatstone's Bridge. — The combination of resistances which is more commonly used than any other for a comparison of two of them is known as a Wheatstone's network; and when it is embodied in a piece of apparatus for measuring resistances, it is called a Wheatstone's bridge.

A Wheatstone's network consists of six conductors connecting four points; in one of the conductors is a source of E. M. F., and in another branch is a galvanometer or other sensitive detector of current.

In the diagram, Figure 343, A and D are maintained at a fixed potential by the battery B'. Then the fall of potential from A to D is the same whether one considers the path ABD or the path ACD; and if a point, B, is chosen on the one path, another one, C, may always be found on the other such that the fall of potential from A to B is the same as from A to C; in other words, B and C have the same potential, and if these points are connected through a galvanometer, no current will flow through it. If C has the same potential as B, by the principle of the last article the ratio of the fall of potential in AB to that in BD is equal to R_1/R_3, and the ratio of the fall in AC to that in CD is equal to R_2/R_4. But these ratios are equal to each other; hence

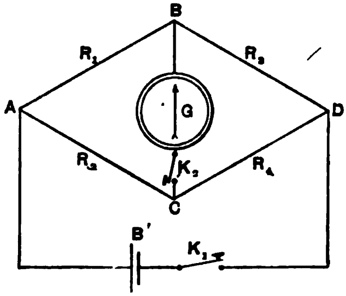


Fig. 343

$$\frac{R_1}{R_3} = \frac{R_2}{R_4}. \tag{100}$$

In practice three of these resistances are fixed and the adjustment for a balance is made by varying the fourth. It is necessary to know only the ratio $\frac{R_2}{R_4}$, for example; the equation (100) gives the relation between R_1 and R_3.

559. Cells in Series. — If n similar voltaic cells, with an electromotive force E and internal resistance r for each cell, are joined in series by connecting the negative of the first with the positive of the second, and so on through the series; and if R is the resistance external to the battery, by the application of Ohm's law the current is

$$I = \frac{nE}{R + nr}. \tag{101}$$

If R is small in comparison with r, $I = E/r$ nearly, or the current is no greater than could be obtained from a single cell; but if R is large in comparison with r, then the current with n cells in series is nearly n times as large as one cell would give with the same external resistance. The internal resistance of a storage cell is very low, and the short-circuiting of a single cell gives as large a current as the short-circuiting of a large number of cells in series.

560. Cells in Parallel. — A battery is said to be connected *in parallel*, or *in multiple*, when all the positive electrodes are joined together for the positive terminal, and all the negative electrodes for the negative. The object aimed at is the reduction of the internal resistance, except that storage cells are joined in parallel when it is desired to draw from them a larger current than the normal discharge current for a single cell. With several cells in parallel, the current in the external circuit is divided among them.

When n similar cells are connected in parallel, the E. M. F. is the same as for a single cell, but there are n internal paths of equal resistance, and the internal resistance is r/n. Hence

$$I = \frac{E}{R + r/n}. \qquad (102)$$

Problems

1. Three Daniell cells are joined in series; the E. M. F. of each cell is 1.1 volts and the internal resistance 2 ohms. If the external resistance is 5 ohms, find the current.

2. Two Leclanché cells are joined in parallel; each has an E. M. F. of 1.5 volts and an internal resistance of 4 ohms. If the external resistance consists of two parallel conductors of 2 and 3 ohms respectively, find the current through each branch.

3. A cell whose E. M. F. is 2 volts gave a current of $\frac{1}{4}$ ampere through an external resistance of 3 ohms. What was the internal resistance of the cell?

4. Four similar cells, each having an E. M. F. of 1.5 volts, are joined in series through a resistance and give a current of one ampere; and when joined in parallel through the same resistance, the current is one third less. What is the internal resistance of each cell?

5. Three wires are connected in parallel; their resistances are 20, 30. and 60 ohms. Find their combined resistance in parallel.

6. The resistance between two points A and B of a circuit is 25 ohms; when another wire is connected in parallel between A and B, the resistance becomes 20 ohms. Find the resistance of the second wire.

7. A current of $\frac{1}{4}$ ampere flows through an incandescent lamp under a pressure of 220 volts. What is the resistance of the lamp?

8. What is the resistance of a set of coils carrying a current of 10 amperes when the fall of potential between the terminals of the coils is 450 volts?

9. If the ratio of two of the resistances forming a Wheatstone's bridge is 10, and the resistance of the third branch is 20 ohms, what is the resistance of the fourth branch for a balance?

10. Two voltaic cells are joined in series with a given resistance and produce a current of 3 amperes; one of the cells is then reversed and the current falls to 1 ampere. Find the ratio of the electromotive forces of the two cells.

IV. THERMAL EFFECTS OF A CURRENT

561. Conversion of Electric Energy into Heat. — It is now a familiar fact that electric energy is readily converted into other forms. If an electric current encounters a back E.M.F. anywhere in the circuit, work is done by the passage of the current against this opposing E.M.F. Such is the case in most examples of electrolysis and in charging a storage battery. All the energy of an electric current not so converted, or stored in some form of stress as potential energy, is dissipated as heat. Heat is generated wherever the circuit offers resistance to the current. In a simple circuit containing no devices for transforming and storing energy, all of it is converted into heat.

The heating of a carbon or a tungsten filament to incandescence by the passage of an electric current is a fact of

everyday observation. Electric cars are heated by currents passing through resistance coils.

The heat evolved by dissolving 33 gm. of zinc in sulphuric acid Favre found to be 18,682 calories. When the same weight of zinc was consumed in a voltaic cell, the heat evolved in the entire circuit was 18,674 calories. The operations were conducted in both cases in a large calorimeter. The two quantities are nearly identical, or the heat generated is the same whether the solution of the zinc is associated with an electric current or not. When a definite amount of chemical action takes place in a battery and no work is done, the distribution of the heat evolved is altered, but not its amount.

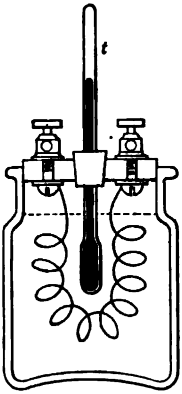


Fig. 344

562. Laws of the Generation of Heat. — The laws of the generation of heat in an electric circuit were discovered experimentally by Joule and Lenz. The latter experimented with a simple calorimeter like that of Figure 344. A thin platinum wire, joined to two stout conductors, was inclosed in a wide-mouthed bottle containing alcohol. A thermometer *t* passed through a hole in the stopper. The resistance of the wire was known, and the observations consisted in measuring the current and noting the rise of temperature. Joule found that the number of units of heat generated in a conductor is proportional

a. To its resistance;
b. To the square of the current strength;
c. To the length of time the current flows.

563. The Heat Equivalent of a Current. — When Q units of electricity are transferred through a potential difference of V units, both in *c. g. s.* measure, the work done is QV ergs, since the potential difference between two points is defined as the work required to transport one unit of electricity from the one point to the other.

If I' is the current in *c.g.s.* units, $Q = I't$. Hence the work

$$W = I'Vt \text{ ergs.}$$

But $W = JH$ (§ 437). Substituting for W in the last equation and solving for H,

$$H = \frac{I'Vt}{J} = \frac{I'Vt}{4.186 \times 10^7} \text{ calories.}$$

If, however, the current is measured in amperes I and the potential difference in volts E, then since $I' = I \times 10^{-1}$ and $V = E \times 10^8$ (§ 584), it follows that

$$I'Vt = IEt \times 10^7 \text{ ergs,}$$

and
$$H = \frac{IEt \times 10^7}{4.186 \times 10^7} = \frac{IEt}{4.186} \text{ calories.} \qquad (103)$$

But by Ohm's law $E = RI$. Then

$$H = \frac{I^2Rt}{4.186} = 0.239 \times I^2Rt \text{ calories.} \qquad (104)$$

Obviously equation (103) may be written

$$HJ = IEt \times 10^7 \text{ ergs} = IEt \text{ joules.}$$

This relation has been used by Griffiths and by Barnes as the basis of a very accurate electrical method of determining J. By passing a current through a coil of platinum wire in a water calorimeter and measuring I, E, and H for an observed time of t seconds, all the quantities in the equation are known except J, which may then be calculated.

564. Counter E. M. F. in a Circuit. — The total activity, or rate at which a generator is supplying energy to a circuit, is represented in part by the heat evolved in accordance with Joule's law, and in part by work done, such as chemical decomposition and storage in electrolysis, the mechanical

work of a motor, etc. In every case of doing work the energy absorbed by it is proportional to the current strength instead of its square. But the whole energy expended per second by the generator is the product of the current strength and the applied E. M. F. We may therefore write for the whole energy transformed in t seconds,

$$IEt = I^2Rt + AIt.$$

The first term of the second member of this equation is the waste in heat; the second, the work done; A is a constant or proportionality factor. Dividing through by It and solving for I,

$$I = \frac{E - A}{R}.$$

It is obvious from the form of this equation that the constant A is of the nature of an E. M. F. Since it has the negative sign, it is a counter E. M. F. The effective E. M. F. producing the current in accordance with Ohm's law is the applied E. M. F. less the counter E. M. F. This counter E. M. F. is necessarily present in every case in which work is done by an electric current.

565. Expression for the Work Done. — If the counter E. M. F. be represented by E', the equation for the current by Ohm's law is

$$I = \frac{E - E'}{R}. \tag{105}$$

Therefore the heat waste in watts becomes

$$I^2R = I(E - E') = IE - IE'.$$

Now IE is the total activity in the portion of the circuit considered, that is, the whole energy applied to it per second. The heat generated per second in this same portion of the circuit of resistance R is less than the energy applied by IE' watts. Hence, the energy spent per second *in doing*

work is the product of the current strength and the counter E. M. F.

The ratio of the work done to the heat waste is

$$\frac{IE'}{I(E-E')}=\frac{E'}{E-E'}.$$

With a given value of E, the value of this ratio increases with E'; that is, the *efficiency* with which electric energy is converted into work increases with the counter E. M. F.

566. Applications of Electric Heating. — Of the many applications of heating by an electric current the following are some of the more important:

1. *Electric cautery.* A thin platinum wire heated to incandescence is employed in surgery instead of a knife. Platinum is used because it is infusible except at a high temperature, and it is not corrosive.

2. *Safety fuses.* Advantage is taken of the low temperature of fusion of some alloys, in which lead is a large constituent, in making safety fuses to open a circuit automatically whenever the current becomes excessive. Safety fuses should be mounted on noncombustible bases or inclosed in a protecting tube.

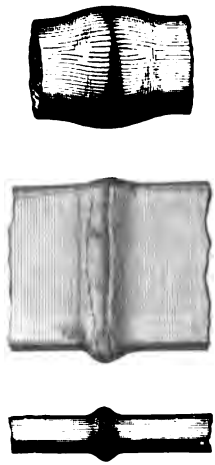


Fig. 345

3. *Electric heating.* Electric street cars are sometimes heated by a current through suitable iron wire resistances embedded in cement, asbestos, or enamel. Similar devices for cooking have now become articles of commerce. Small furnaces for fusing, vulcanizing, and enameling in the operations of dentistry are in common use. Large furnaces for melting iron and the reduction of iron ores by electric heat are in use to some extent. They are also employed in many chemical operations requiring high temperatures.

4. *Electric welding.* If the abutting ends of two rods or bars are pressed together while a large current passes through them, enough heat is generated at the junction, where the resistance is greatest, to soften and weld them together. The rod becomes uniform when the weld is complete and the heat is no longer localized. This method has been perfected by Elihu Thomson. Figure 345 shows three small welded joints.

V. ELECTRIC LIGHTING

567. The Carbon Arc. — In 1800 Sir Humphry Davy discovered that when two pieces of charcoal, connected by suitable conductors to a powerful voltaic battery, were brought into contact and then separated a slight distance, brilliant sparks passed between them; but no mention was made of the electric arc until 1808. In 1810 Davy exhibited the arc light at the Royal Institution in London.

With a battery of 2000 simple elements, when the carbons in a horizontal position were drawn apart to a distance of several inches, the carbon was apparently volatilized, and the current was conducted across in the form of a curved flame or *arc*. Hence the name *electric arc* was given to this form of electric lighting.

The white-hot charcoal electrodes of Davy burned away very rapidly unless they were inclosed in a vacuum. It was not till 1844 that Foucault surmounted the difficulty by the use of dense carbon from a gas retort instead of wood charcoal.

When the carbon points are separated, the heat due to the current volatilizes some of the carbon and ionizes it, and this ionized carbon vapor conducts the current across. The dazzling light is emitted chiefly by the vividly hot carbon electrodes, and especially by the positive one. In it is formed a small cavity by the transport of carbon across to the negative. Violle has estimated the temperature of this depression or *crater* at 3500 C. The positive carbon rod wastes away about twice as fast as the negative one.

Duddell has found a thermal electromotive force between hot carbon and carbon vapor, and directed from the latter to the former. At the positive electrode there is therefore a back E. M. F. against which the current flows; the current thus gives up energy there, and the only form it can take is heat. The reverse is true at the negative electrode. The principal resistance to the passage of the current is offered by the layer of carbon vapor near the electrodes. The counter E. M. F. at the positive accounts for the higher initial temperature on this side; a denser

layer of carbon vapor is thus formed there, which in turn increases the resistance with the generation of more heat than at the negative.

The efficiency of an electric lamp is defined as the ratio of the electrical power expended to the candle power emitted. The efficiency of the carbon arc is about 1 watt per candle power.

568. The Inclosed Arc. — When the electric arc is maintained between rods of hard retort carbon in the open air, the carbon burns away quite rapidly. The potential difference between the carbons is then from 45 to 55 volts for a 10-ampere current.

In the "inclosed arcs" the lower carbon and a portion of the upper one are inclosed in a small globe, which is air-tight at the bottom, but allows the upper carbon to slip through a check-valve at the top. Soon after the arc is formed, the oxygen is absorbed and the arc is thereafter inclosed in an atmosphere of nitrogen and carbon monoxide. The inclosed arc is longer than the open arc, and the potential difference required to maintain it is about 80 volts, but the current for the same consumption of energy is smaller than the open arc requires. The carbons for the inclosed arc last about ten times as long as in the open air.

569. Other Arc Lights. — Other arc lamps are now in commercial use in which the light comes chiefly from the incandescent stream between the electrodes. They have a higher efficiency than the carbon arc. In the "metallic arc" powdered *magnetite* in an iron tube is used for one electrode and a block of copper for the other. The arc flame is very white and brilliant; the light comes from the luminous iron vapor.

"Flaming arcs" are made by the use of a positive electrode impregnated with salts of calcium. The light from the flaming arc is yellow and is adapted to outdoor illumination only.

570. The Incandescent Lamp. — In the incandescent lamp the heat is due to the simple resistance of a thin conducting filament inclosed in an exhausted glass bulb. The terminals of the filament are connected through the glass by means of two short pieces of platinum wire. Platinum is used for this purpose because its coefficient of expansion is about the same as that of glass; and so, when the lamp becomes hot in use, it neither leaks air around the wires nor cracks.

The carbon filament is made usually from cellulose obtained from cotton. After preliminary treatment it is carbonized by raising to a cherry-red heat out of contact with the air. It is then surrounded by an atmosphere of rarefied hydrocarbon vapor, and is heated white hot by an electric current. The heat decomposes the vapor, and the carbon residue is deposited in a dense form on the filament. By this treatment the filament acquires a hard, steel-gray surface and greater uniformity.

The temperature to which a carbon filament can be raised is limited by the tendency of the carbon to disintegrate at high temperatures. This disintegration rapidly reduces the thickness of the filament and blackens the glass bulb.

In recent years filaments have been made of the rare metals osmium, tantalum, tungsten, and some other materials. The tungsten lamp is rapidly displacing the carbon lamp because of its higher efficiency, in spite of the fact that it is much more fragile. Its efficiency is as high as from 1 to 1.25 watts per candle power.

The ordinary commercial unit for the carbon filament is the 16-candle-power incandescent lamp. On a 110-volt circuit it takes about 0.5 ampere. Since the power in watts consumed is EI, this lamp consumes about 55 watts, or 3.5 watts per candle power. The tungsten 25-watt lamp gives 20 candle power, and the 40-watt lamp 32 candle power.

VI. THERMOELECTROMOTIVE FORCE

571. Thermoelectric Junction. — If a circuit be formed by two wires of different metals joined together at their ends, no current will flow through it so long as both junctions are at the same temperature. Any electromotive forces at the two junctions are then equal and in opposite directions around the circuit. Seebeck found about 1822 that if one junction is at a higher temperature than the other, there is in general an electromotive force in the circuit and a current will flow.

If a copper wire and an iron wire be twisted together and their free ends be connected to a sensitive galvanometer, and if the twisted junction be warmed to a higher temperature than the rest of the circuit, a current will flow across the warmed junction from copper to iron. Such a current is called a *thermoelectric current*, and the electromotive force at the junction of the wires, *thermoelectromotive force.*

The dissimilar substances composing a *thermoelectric junction* or *pair* may be two metals, a metal and a liquid, two liquids, or even two pieces of the same metal at different temperatures or in different physical states.

572. The Neutral Temperature. — Assume two copper wires connected to a suitable galvanometer and their free ends joined by an iron wire. If now one junction of the copper and iron be kept at a constant temperature, while that of the other is gradually raised, the current in the circuit will increase up to a certain temperature of the hot junction. This temperature is called the *neutral temperature* or *neutral point.*

If the hot junction be heated still more, the current will decrease and, finally, it will become zero when the temperature of the hot junction is as much above the neutral point as that of the cold one is below it. If the temperature of the hot junction be raised still higher, there will again be a current, but its direction will be reversed.

573. The Thermopile. — The E. M. F. of a single thermal couple is very small; to get a larger E. M. F. a number of similar couples may be joined in series. With *n* couples in series the potential difference between the extreme terminals is *n* times that of a single couple. Figure 346 shows the manner of connecting in series. If the bars are *A* antimony and *B* bismuth, then heating the junctions *c*, *c*, *c* will cause a current to flow through the circuit in the direction of the arrow; but if these junctions are cooled, or the alternate ones, *d*, *d*, heated, the current will flow in the other direction.

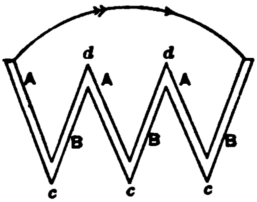


Fig. 346

When a number of bars of antimony and bismuth are soldered together in this manner, and are closely packed in the form of a cube, with insulating material between adjacent bars, so that opposite faces of the cube form alternate junctions, the instrument is called a *thermopile* (Fig. 347). When the face of such a pile is blackened with lampblack and is provided with a reflecting cone, the instrument becomes a sensitive detector of radiation (§ 427).

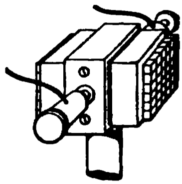


Fig. 347

574. The Peltier Effect. — In 1834 Peltier discovered the phenomenon which bears his name; it is the converse of Seebeck's discovery. If a bismuth-antimony junction be heated,

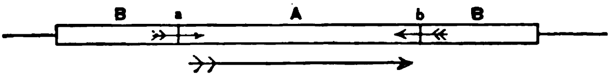


Fig. 348

the current flows from the former to the latter. Peltier discovered that if a current from an external source be sent through such a compound bar from bismuth *B* to antimony *A* (Fig. 348), the junction will be cooled; but if it be sent the other way, the junction will be heated.

The long arrow shows the direction of the current sent through the bar; the small arrows at *a* and *b* indicate the direction of the thermoelectromotive forces at the junctions. At *a* the thermal E. M. F. is in the direction in which the current is flowing. Hence at this junction work is done on the current, and the heat of the metals is converted into the energy of the current. At *b* the thermal E. M. F. opposes the current, which therefore does work on the junction and heats it.

The thermal effect of a current at a junction of dissimilar substances differs greatly from the thermal effect due to simple ohmic resistance. The Peltier effect is *reversible*, the current heating or cooling the junction according to its direction, and the quantity of heat evolved or absorbed varies simply as the current strength; the heat due to resistance is independent of the direction of the current, and is proportional to the square of the current strength.

575. Experiment to demonstrate the Peltier Effect. — Connect a Leclanché cell, *B′*, with a thermopile and a sensitive galvanometer, as in Figure 349. *S* is a two-point switch; when it is turned in the direction of the full line, the circuit through the cell and the thermopile is closed and the galvanometer circuit is open. When it stands in the direction of the dotted line, the cell is cut off and the thermopile is connected to the galvanometer. To show that the current from the thermopile *P* is opposite in direction to the current sent through it by the voltaic cell, insert in the circuit of the galvanometer at *T* a copper-iron couple. With the switch at *b*, the current produced by heating this junction flows from *Cu* to *Fe*, and the direction of the galvanometer deflection may be noted. Turn the switch for a moment to *a* and then back again to *b*. The galvanometer will show a current coming from the thermopile, and the direction of the deflection will be the same as when the junction *T* was warmed. Hence *B* must be the positive and *A* the negative electrode of the thermopile as a generator. But the current from

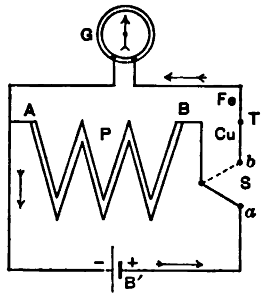


Fig. 349

the cell enters the pile at B and leaves it at A. The thermal effect produced by the current through the pile is such as to generate a counter E. M. F.

576. Thermoelectromotive Force between Metals and Liquids. —The thermoelectromotive forces having their seat at metal-liquid contacts have special interest because of their relation to the temperature coefficient of voltaic cells. These electromotive forces are larger than most of those between metals. Thus, the thermoelectromotive force of Zn-$ZnSO_4$ is 0.00076 volt per degree for a mean temperature of 18.5°; that of Cu-$CuSO_4$ is 0.00069 for about the same mean temperature. In microvolts (millionths of a volt) these are 760 and 690 respectively. Since the metal is in both cases positive to the solution, and there is no appreciable E. M. F. at the contact of the two liquids, the temperature change per degree in the E. M. F. of a Daniell cell is the difference between the thermoelectromotive forces at the zinc and copper electrodes, or 0.00007 volt per degree. It is, moreover, negative because the thermal E. M. F. on the zinc side is greater than on the copper side. A rise of temperature of one degree on the zinc side of the cell lowers the E. M. F. 0.00076 volt; the same rise of temperature on the copper side of the cell raises the E. M. F. 0.00069 volt. The same rise of temperature of the whole cell therefore lowers the E. M. F. by the difference of the two thermoelectromotive forces, or by 0.00007 volt. This conclusion has been fully verified by experiment.

This method of analysis of the temperature coefficient has been applied to other cells, such as the Clark standard cell and the Weston normal cell; the results in every case show that the temperature coefficient is determined by the thermoelectromotive forces at the contacts of the dissimilar substances in the cell, whenever it is not complicated by the solution and recrystallization of salts.

The Peltier phenomenon applies to junctions between solids and liquids. When a Daniell cell furnishes a current,

heat is absorbed at the positive or copper electrode and is generated at the zinc electrode, because the thermoelectromotive force on both sides is directed from the liquid to the metal. Research has shown that the observed difference of temperature between the two sides of a suitably designed Daniell cell, due to the flow of a known quantity of electricity through it, conforms to the calculated value.

With metal-liquid junctions there may be an E. M. F. in the circuit without any differences of temperature, because the thermoelectromotive force is a function of the concentration of the solution as well as of the temperature.

Problems

1. The electrodes of a voltaic cell are joined by two wires alike in every respect, except that one is twice as long as the other. What are the relative quantities of heat generated in the two?

2. The E. M. F. of a battery is 20 volts and its internal resistance 2 ohms. The potential difference between its poles when connected by a wire A is 16 volts; it falls to 14 volts when A is replaced by another wire B. Calculate the number of calories of heat generated in the external circuit in 3 min. in the two cases.

3. A current of 10 amperes flows through a resistance of 2 ohms for 14 sec. Find the number of calories of heat generated.

4. What current would have to flow for an hour through a resistance of 20 ohms to produce enough heat to raise the temperature of a kgm. of water from the freezing point to the boiling point?

5. 500 incandescent lamps in parallel are supplied with one half ampere each at a potential difference of 110 volts between lamp terminals. The drop of potential in the line is 2.2 volts. What is the resistance of the line and how much power is lost in it?

6. A battery has an E. M. F. of 8.5 volts; the total resistance in the circuit is 20 ohms, including an electrolytic cell. The heat generated per second in a 5.12 ohm coil included in the circuit is 0.12 calorie. What is the counter E. M. F. of the electrolytic cell?

CHAPTER XIX

ELECTROMAGNETISM

I. MAGNETIC RELATIONS OF A CURRENT

577. Oersted's Discovery. — The discovery by Oersted at Copenhagen in 1819 was one of prime importance, for he was the first to find any connection between electricity and magnetism. He observed that when a magnetic needle is brought near a long straight wire conveying a current, the needle tends to set itself at right angles to the length of the wire; also that the direction in which the needle turns depends on the direction of the current through the conductor. The experiment of Oersted shows that the region around a wire conveying an electric current has magnetic properties, that is, it is a magnetic field. At this point the analogy between an electric current and a stream of water flowing through a pipe fails, for such a stream produces no effect in the region surrounding the pipe.

Assume a current flowing through the conductor above the needle *NS* from north to south, as indicated in Figure 350. The north pole will turn toward the east. If the current be reversed, or if it be placed under the needle while still flowing in the same direction, the north pole will turn toward the west.

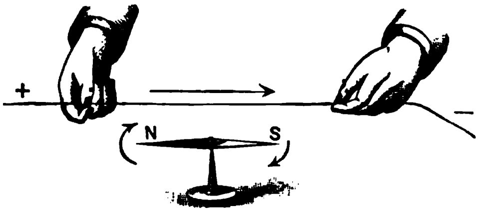


Fig. 350

If the wire be carried around the needle in a rectangular loop (Fig. 351), both branches of it will contribute to the force of deflection, and the north-seeking pole at the left will turn toward the east.

578. Direction of Deflection with Respect to the Current. — All the movements of a magnetic needle under the influence of a current may be summed up in one rule:

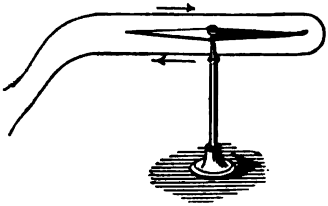


Fig. 351

Stretch out the right hand in the direction of the wire, with the palm turned toward the magnetic needle, and with the current flowing in the direction of the extended fingers; the outstretched thumb will point in the direction in which the north pole is deflected.

579. Magnetic Field about a Wire. — A little consideration will show that if the current flows in the direction of the long arrow (Fig. 352), the resulting magnetic force is in the direction of the small arrows around the circle; conversely, if the current flows around the circle clockwise, the positive direction of the magnetic force, or the direction in which a north pole is urged, is along the axis of the circle away from the observer. If the fingers of the closed right hand represent the circle with the current flowing around in the direction of the finger tips, the outstretched thumb points in the direction of the lines of force, and conversely (§ 580). The lines of force due to a current are, therefore, concentric circles about the conductor as a center.

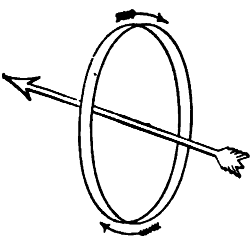


Fig. 352

Such circular lines of force may be shown by a fairly strong current through a vertical wire passing through a hole in a horizontal glass plate,

on which are evenly sifted fine iron filings. When the plate is gently tapped, the filings are left free to arrange themselves in circles, indicating the lines of magnetic force around the wire. Figure 353 is reproduced from a photograph of such circular lines. They show that the ether about the current is under stress, and therefore possesses potential energy. It is rather more important to direct the attention to this magnetic stress in the ether than to what goes on within the conductor itself.

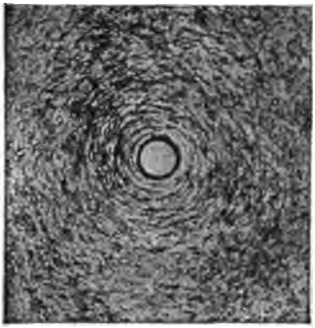


Fig. 353

580. Magnetic Properties of a Circular Conductor. — Bend a copper wire into the form shown in Figure 354, and suspend it by a long untwisted thread so that the ends dip into the mercury cups shown in section at the bottom. When a current is sent through the suspended wire, a magnet pole near the circular conductor will cause the latter to turn around a vertical axis and take up a position with its plane at right angles to the axis of the magnet.

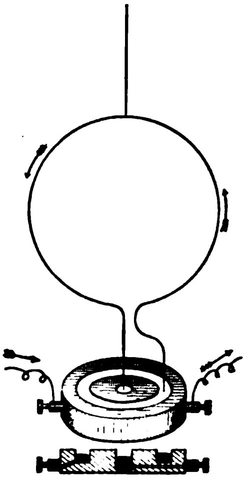


Fig. 354

This experiment shows that a current in the form of a loop acts like a magnetic shell or disk. The lines of force about the circular conductor pass through it and come around from one face to the other through the air outside the loop. An electric circuit is in every case equivalent to a magnetic shell whose contour coincides with the circuit. The clŏsed circuit and the magnetic shell have in their vicinity similar magnetic fields.

The north-seeking side of the loop is the one from which the magnetic lines issue; to an observer looking toward this

side, the current flows around the loop counter-clockwise (Fig. 355). The direction of the current and that of the lines of force are related to each other as the direction of rotation and of translation in a right-handed screw.

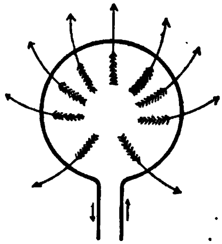


Fig. 355

581. Magnetic Intensity at the Center of a Circular Coil. — The intensity of the magnetic field at any point is the force in dynes acting on a unit magnetic pole (§ 459) placed at the point. Faraday showed that the magnetic intensity produced by a current is proportional to the current strength; Biot and Savart demonstrated experimentally that for a current of indefinite extent it is inversely proportional to the distance between the conductor and the point; Laplace proved that this latter result follows from the law of inverse squares as applied to the mutual action between an element of the conductor and unit pole at the point. Hence the intensity due to the current in an *element* l of the conductor, at a point P on a perpendicular from the element, is

$$\mathcal{F} = k\frac{Il}{r^2},$$

where r is the distance between the current element and the point P, k is the constant or proportionality factor, and I the strength of the current.

If the point P is at the center of a circle of radius r and a current I is flowing around the circle, then the magnetic intensity at the center due to the current in the entire circumference is

$$\mathcal{F} = k\frac{I\Sigma l}{r^2} = k\frac{2\pi r I}{r^2} = k\frac{2\pi I}{r}.$$

If now the unit current is so defined as to make k equal to unity, then

$$\mathcal{F} = \frac{2\pi I}{r}. \tag{106}$$

582. The Electromagnetic Unit of Current. — The electromagnetic system of electrical units in common use is based on the magnetic effects of a current. The starting point is the magnetic intensity due to a conductor conveying a current.

If an element of a conductor one centimeter long be bent into an arc of one centimeter radius, the current through it will have unit strength when it exerts a force of one dyne on a unit pole at the center of the arc.

This definition is equivalent to making the constant k in the last article equal to unity. If the field due to unit current in unit length of the conductor is unity, the field due to the same current through the whole circumference will be 2π; and if the current is I units, the field will be $2\pi I$. If, further, the radius is not unity, but r, the circumference will be $2\pi r$, and then

$$\mathcal{F} = \frac{2\pi r I}{r^2} = \frac{2\pi I}{r}.$$

The *ampere* is one tenth of this *c. g. s.* electromagnetic unit of current. The unit of quantity in the electromagnetic system is the quantity which passes any cross section of the conductor in one second when the current through it is one *c. g. s.* unit. The practical unit of quantity is the *coulomb;* it corresponds with the ampere, and is one tenth of the *c. g. s.* unit of quantity.

583. Electromagnetic Units. — It will be convenient for reference to bring together the several electrical units expressed in electromagnetic measure in the *c. g. s.* system.

Unit Strength of Current. A current has unit strength when a length of one centimeter of its circuit, bent into an arc of one centimeter radius, exerts a force of one dyne on a unit magnetic pole at its center.

Unit Quantity. Unit quantity is the quantity conveyed by unit current in one second.

Unit Potential Difference. Unit potential difference exists between two points when the transfer of unit quantity from one point to the other requires the expenditure of one erg of work.

Unit Resistance. A conductor offers unit resistance when unit potential difference between its ends causes unit current to flow through it.

Unit Capacity. A conductor has unit capacity when unit quantity charges it to unit potential.

584. Practical Electromagnetic Units. — Several of the *c.g.s.* electromagnetic units are inconveniently small and others are inconveniently large for practical purposes. Hence the following multiples and submultiples of them have been universally adopted as the *practical units:*

Current. The *ampere*, equal to 10^{-1} *c.g.s.* unit; it is represented by the current which will deposit silver from silver nitrate solution at the rate of 0.0011182 gm. per second.

Quantity. The *coulomb*, equal to 10^{-1} *c.g.s.* unit of quantity; it is the quantity conveyed by a current of one ampere in one second.

Electromotive Force. The *volt*, equal to 10^8 *c.g.s.* units; it is 10,000/10,184 of the E.M.F. of a Weston normal cell at 20°.

Resistance. The *ohm*, equal to 10^9 *c.g.s.* units; a volt produces an ampere through a resistance of one ohm; practically the ohm is represented by the resistance at 0° of a uniform thread of mercury 106.3 cm. in length and 14.4521 gm. mass.

Capacity. The *farad*, equal to 10^{-9} *c.g.s.* unit; it is the capacity of a condenser which is charged to a potential difference of one volt by one coulomb. The microfarad, chiefly used in practice, is one millionth of a farad, or 10^{-15} *c.g.s.* unit.

Work. The *joule*, equal to 10^7 ergs; it is represented by the energy expended per second by one ampere under a pressure of one volt.

Power. The *watt*, equal to 10^7 ergs per second; it is equivalent to the power of a current of one ampere flowing under a pressure of one volt, or to one joule per second; it is very approximately $\frac{1}{746}$ of a horse power.

Induction. The *henry*, equal to 10^9 *c.g.s.* units; it is the induction in a circuit when the electromotive force induced in this circuit is one volt, while the inducing current varies at the rate of one ampere per second (§ 623).

The prefixes *kilo-* and *milli-* combined with any of the preceding units signify a *thousand* and a *thousandth* respectively. Thus, a kilowatt is a thousand watts, and a millivolt is a thousandth of a volt. The prefixes *mega-* and *micro-* signify a *million* and a *millionth* respectively. Thus, a megohm is a million ohms, and a microfarad is a millionth of a farad.

II. GALVANOMETERS

585. Types of Galvanometers. — When currents are compared by means of their magnetic effects, the instrument used for the purpose is called a *galvanometer*.

The three types of galvanometers most in common use are as follows: (1) those in which the current through a fixed coil of wire causes a deflection of a suspended magnetic needle, usually at the center of the coil; (2) those in which the coil itself is movable around a vertical axis between the poles of a fixed magnet; (3) these two types are applicable to direct currents only; for both direct and alternating currents another kind is employed, in which both the fixed and the movable parts are coils; these are known as *electrodynamometers*.

586. Nobili's Astatic Pair. — Galvanometers of type (1) are commonly used to detect small currents. The first requisite in such instruments is sensitiveness, that is, a very small current must produce an observable deflection of the needle. For this purpose the controlling couple of the earth's field on the movable magnetic system must be reduced. This may be done by means of a weak compensating magnet, placed either above or below the movable magnetic needle, with its north-seeking pole turned toward the north. The field produced by it is then opposed to the earth's field.

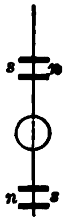


Fig. 356

Nobili's astatic pair illustrates the other method in common use. It consists of a pair of needles, or two groups of needles, mounted in the same vertical plane, but with their similar poles turned in opposite directions (Fig. 356). If the two needles have equal magnetic moments, the resultant action of the earth's field on the system is zero. Since such a system has no directive tendency, it is called *astatic*, because it will remain at rest in any azimuth. In practice the astatic system is never exactly balanced magnetically, and the earth's field always has some directive influence.

If both needles are surrounded with coils so connected that the current flows around them in opposite directions, the two forces of deflection turn the system in the same direction, while the opposing controlling force is reduced to a small value.

587. The Astatic Mirror Galvanometer. — In Figure 357 the front coils are swung open to expose to view the astatic system. It consists of minute pieces of magnetized watch-spring at the centers of the coils above and below. In very sensitive instruments like this one a small mirror is attached to the movable system; the instrument is then called a *mirror galvanometer*. In the figure the mirror is midway between the two sets of small magnets. Sometimes a beam of light from a lamp is reflected from the small mirror back to a scale, and sometimes the light from a scale is reflected back to a telescope, by means of which the deflections are read. In either case the beam of light becomes a long pointer without weight.

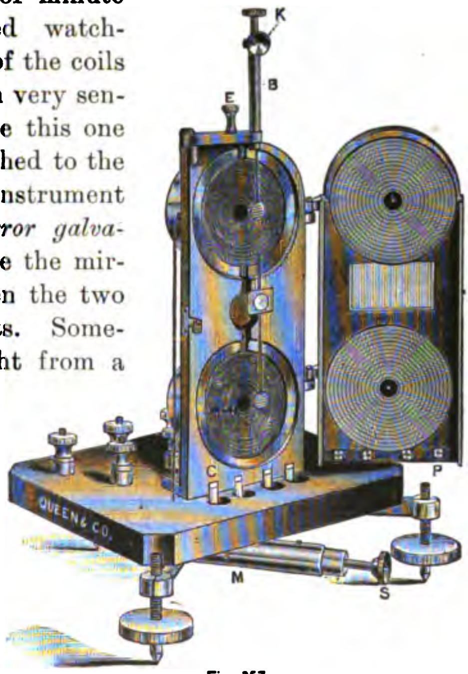


Fig. 357

588. The D'Arsonval Galvanometer. — It is immaterial from a magnetic point of view whether the magnet or the coil of a galvanometer is made movable, since the action between them is reciprocal. In the D'Arsonval galvanometer (Fig. 358) a coil, suspended by a fine wire, swings between

the poles of a strong permanent magnet. The current is led in by the suspending wire and out by the wire or spiral spring connecting the coil to the bottom. The great advantage of this type of galvanometer is that it has a strong field of its own, which is only slightly affected by the earth's magnetism or by iron or other magnetic material in its neighborhood. A small mirror for reflecting a beam of light is attached to the coil. Inside the coil is a soft iron tube supported from the back. This has the effect of strengthening the narrow magnetic field in which the coil swings.

Fig. 358

589. Potential Galvanometers. — Galvanometers designed to determine the potential difference between two points of a circuit must be of high resistance. If they are graduated to read in volts, they are called *voltmeters*. Any sensitive galvanometer may be used as a voltmeter by adding a sufficiently large resistance in series with it. Unless the resistance of the voltmeter is high, the application of its terminals to two points of a circuit A and B (Fig. 359), so as to put it in parallel with a resistance s through which a current is flowing, will diminish the potential difference which it is desired to measure.

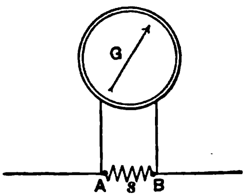


Fig. 359

For direct currents the most convenient portable voltmeter is made on the principle of the D'Arsonval galvanometer. The inside of the instrument is shown in Figure 360. A portion of the magnet and of one

pole piece is cut away to show the coil and the springs. The current is led in by one spiral and out by the other. Attached to the coil is a light aluminum pointer, which moves over a scale on which the voltage is read directly.

Fig. 360

A similar instrument, called an *ammeter*, is designed to measure currents in amperes. The resistance of a voltmeter should be so high that it will take the smallest operating current; the resistance of an ammeter should be as small as possible, so that it will not increase the resistance of the circuit in which it is placed.

III. ELECTRODYNAMICS

590. Magnetic Field about Parallel Currents. — The term *electrodynamics* is applied to that part of the science of electricity which is concerned with the force exerted by one current on another. The reciprocal action between conductors conveying currents was discovered by Ampère in 1821, shortly after Oersted's discovery of the reciprocal action between a current and a magnet.

Every conductor through which a current is flowing is surrounded by a magnetic field, and the magnetic fields of two such conductors react on each other. The reciprocal action between conductors carrying currents is purely magnetic, and may be accounted for by the stresses set up in the surrounding medium.

The magnetic field about a single conductor is composed, as we have seen, of concentric circles; but when the fields of two conductors are in part superposed, the composite magnetic figures will be those due to the resultant of the two sets of forces in the field.

Figure 361 is the field shown by iron filings about two parallel wires passing through the two holes and with the currents flowing in the same direction. In addition to the small displaced circles immediately around the conductors, there are continuous curves inclosing both circuits. These are due to the coalescence of a number of circles belonging to the two currents. The two conductors are drawn together by the tension along these inclosing lines.

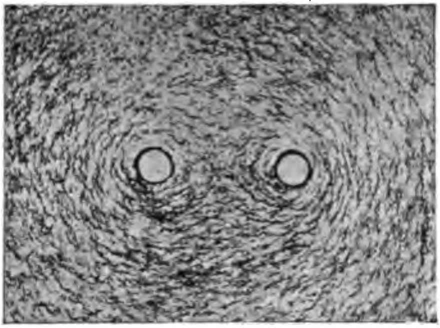


Fig. 361

Figure 362 is the field about two parallel conductors with the currents flowing in opposite directions. Midway between the two wires the lines of force have the same direction in space, and produce a uniform field over a small area. The circles about the two wires are all eccentric, but there are no lines common to the two conductors; the resiliency of these lines, or their tendency to recover from displacement, forces the conductors apart.

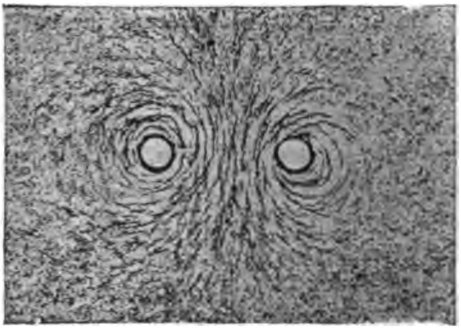


Fig. 362

591. Motion of a Circuit in a Magnetic Field. — The law applying to the electrodynamic action between conductors conveying currents is that *their relative motion is always*

such as to make the flux of magnetic lines around them a maximum. Hence, two circuits tend to move toward coincidence. Each is urged to a position that makes the lines of force common to the two as numerous as possible.

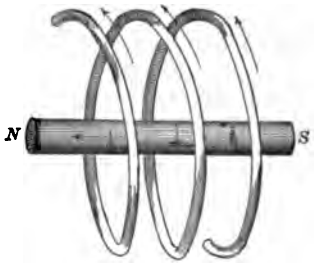


Fig. 363

Similar statements hold with respect to a magnet and a circuit. When a bar magnet and a helix (Fig. 363) come into the relative position where the middle point of the former coincides with the mean plane of the latter, the lines of force of the two are identical in direction through the helix, and the position is one of stable magnetic equilibrium. Hence the law of parallel currents:

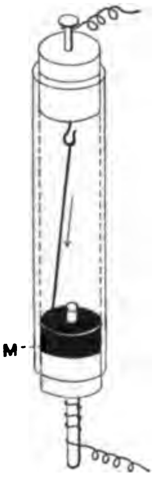


Fig. 364

Parallel conductors conveying currents in the same direction attract each other; if the currents are in opposite directions, they repel each other.

592. Electromagnetic Rotations. — A large number of different devices have been designed for the purpose of producing continuous rotation by the action between a magnet and a circuit. In Figure 364 a copper wire is hung by a hook (better by a ligament of fine copper wires), and the lower end dips into mercury M surrounding the pole of a magnet. If the current flows down through the wire and the upper end of the magnet is a north pole, the bottom of the wire will rotate around the pole clockwise.

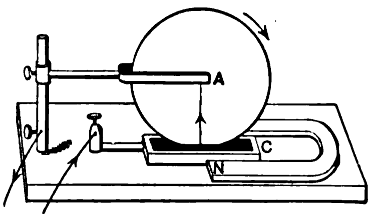


Fig. 365

Barlow's wheel (Fig. 365) is one of the oldest devices to secure con-

tinuous rotation by the action between a current and a magnet. It is therefore a direct current motor without a commutor (§ 635). Contact is made by means of mercury in the trough C, and the action of the magnetic field between the poles of the horseshoe magnet is on the radial current from the mercury to the axis A of the copper wheel.

593. Electrodynamometers.—The electrodynamometer is an instrument designed originally by the German physicist Weber to measure the strength of a current by the electrodynamic action between two coils of wire, one fixed and the other movable about a vertical axis through its own plane. The two coils are set with their magnetic axes at right angles (Fig. 366), and the free coil moves in a direction to make their axes coincide.

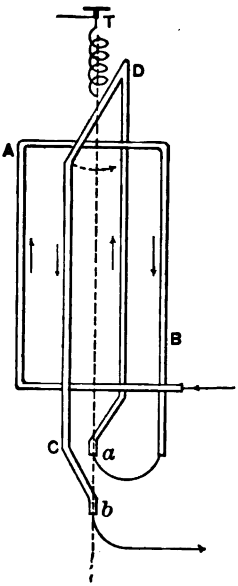


Fig. 366

AB is a single turn of the fixed coil, and *CD* one of the suspended coil. The ends *a* and *b* of the latter dip into mercury cups not shown, and the two coils are in series. The movable coil is suspended by silk threads (or on a point resting in a jewel), and a helix is rigidly connected with it at the top and with a torsion head *T*. The movable conductor is subjected to a system of forces, that is, to a torque, tending to turn it in the direction indicated.

When the coil *CD* is deflected by sending a current through the two coils in series, the torsion head is turned by hand so as to bring the movable coil back to its initial position. The couple due to the magnetic action between the two coils is then in equilibrium with the couple of torsion of the twisted helix. The couple of torsion is proportional to the angle of torsion by Hooke's law. The electrodynamic action between the coils is proportional to the *square* of the current, since

doubling the current doubles it through both coils, doubles the magnetic field of both, and therefore quadruples the force. The square of the current is therefore proportional to the angle through which the helix at the top is twisted to restore the suspended coil to its zero position, or

$$I^2 = A^2 D.$$

Whence $$I = A\sqrt{D}. \tag{107}$$

A is a constant of the instrument depending on the windings and the helix.

Figure 367 is one form of the complete instrument, showing the coils, the helix, and the scale at the top with the pointers, one moving with the suspended coil and the other with the helix. The fixed coil may be considered as furnishing a magnetic field corresponding to that of the permanent magnet in the D'Arsonval galvanometer; but in this instrument the field reverses with the reversal of the current, and therefore the deflection is in the same direction whether the current goes through the instrument in one direction or the other. It may thus be used with alternating or reversing currents as well as with direct ones.

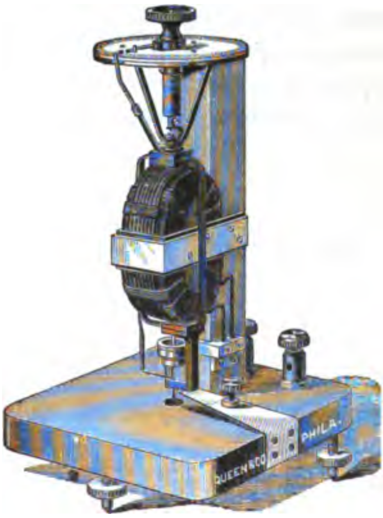


Fig. 367

594. Convection Currents. — Two parallel currents in the same direction attract and two like electric charges repel each other. According to Maxwell, the electrodynamic attraction should exactly equal the electrostatic repulsion when the

electric charges move with the velocity of light. Faraday assumed that a stream of particles (or ions) carrying electric charges of the same sign has a magnetic effect like a current of electricity. Rowland demonstrated the truth of this assumption in 1876; he found that a charged disk, when rapidly rotated, had a feeble field equivalent to a circular current. Conversely, such convection currents are acted on by a magnet. The electric arc behaves like a flexible conductor. It may even be ruptured by the deflecting influence of a powerful magnet. Elihu Thomson has utilized this effect to extinguish an arc started by lightning on an electric lighting circuit.

The titanic whirlpool constituting a sunspot has a magnetic field along its axis. The hypothesis is that the ions composing the whirl are charged, and their rapid rotation in circles is the equivalent of circular currents, with a magnetic axis coinciding with the axis of the whirlpool.

IV. ELECTROMAGNETS

595. Solenoids. — Since a circular current is the equivalent of a plane magnetic shell (§ 580), a helix composed of equal circular currents, all on the same axis, and with their similar faces turned in the same direction, is the equivalent of a cylindrical magnet. Such a system of circular currents constitutes a *solenoid* (Fig. 368). Each turn of the helix may be resolved into a plane circular current ABC (Fig. 369), and a linear current AC perpendicular to the plane of the circle. An entire helix of n turns is therefore equivalent to n circular currents and a linear current along the axis of the helix. If the conductor returns along the axis, as in the figure, the external field is due to the circular elements only.

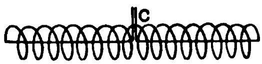


Fig. 368

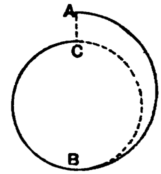


Fig. 369

If such a solenoid is suspended so as to turn freely, it will set its axis in the magnetic meridian when a current is passed through it. It is therefore equivalent to a magnet, and its poles may be determined by the method described in § 579. They will be attracted or repelled by a permanent magnet like those of a magnetic needle.

596. Effect of Introducing Iron. — When an iron bar is introduced into a solenoid through which a current is passing, the iron will be magnetized by the magnetizing force along the axis of the helix. If the iron core is absent, many

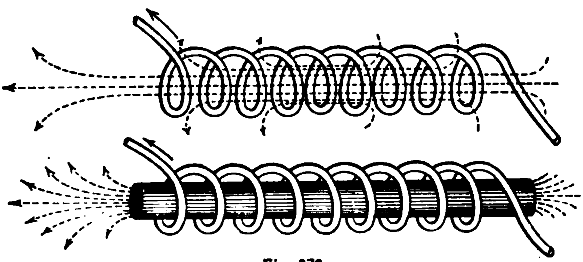


Fig. 370

of the lines of induction leak out at the sides of the helix (Fig. 370). The core not only diminishes the leakage, but adds many more lines to those previously running through the solenoid. Hence the magnetic strength of a helix is indefinitely increased by the iron core.

If the bar be of soft iron, it will exhibit notable magnetism only so long as the current flows through the magnetizing coil. The loss of magnetism is not complete when the current is interrupted; the small amount remaining is known as *residual* magnetism. Temporary magnets produced by magnetic induction within a helix are called *electromagnets*. The north pole of an electromagnet is the one about which the current in the helix appears to circulate counter-clockwise to one looking toward the pole. The circular currents in the helix and the lines of induction in the core are linked together by the right-handed screw relation. *In fact, no*

exception has ever been discovered to the general fact that with every electric current there are always linked lines of induction in this same relation.

Fig. 371

Iron filings arranged in circles about a linear conductor may be regarded as flexible magnetized iron winding itself in a helix around the current; conversely, a free flexible conductor carrying a current winds itself around a straight bar magnet. The flexible conductor of Figure 371 may be made of tinsel cord or braid. Directly the circuit is closed, the conductor winds slowly around the vertical magnet; if the current is reversed, the conductor unwinds and winds up again in the reverse direction.

597. The Horseshoe Magnet. — The most common form of electromagnet is the U-shape or *horseshoe* type (Fig. 372). The windings on the two iron cylinders or *cores* must be in a direction to make the two poles of opposite signs. It is the same as if the two cores were straightened out and the bar wound continuously from end to end. The *armature* (not shown in the figure) consists of a flat bar like the yoke at the other end, and extending across from pole to pole. As a rule, the cores, the yoke, and the armature should form a closed magnetic circuit, that is, one in which the lines of induction are entirely in the iron.

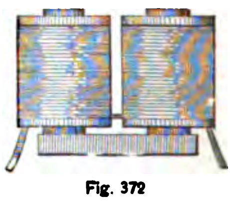


Fig. 372

If a ring be wound continuously with a right-handed helix so as to form a closed circuit, and if connection with the winding be made at two points diametrically opposite

each other (Fig. 373), and a divided current be sent through, there will be a consequent south pole where the current enters and a consequent north pole where it leaves the ring. The poles are consequent because they belong to two magnetic circuits, or to a divided circuit through the iron. Either half of the ring may be regarded as the armature of the other half.

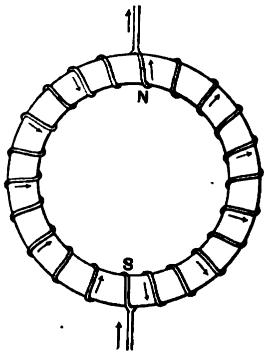


Fig. 373

V. MAGNETIC INDUCTION

598. Magnetic Permeability. — The effect of placing iron in a magnetic field is to increase greatly the number of lines of induction running through the space occupied by the iron. When these lines of induction traverse the iron, it is magnetized. The increase in the number of lines due to the iron may amount to several thousand per square centimeter.

Let $\mathcal{B}$ stand for the magnetic induction, or number of lines per square centimeter through the iron. Then the ratio between $\mathcal{B}$ and $\mathcal{H}$, the magnetizing force, is called the *permeability* of the iron, or

$$\mu = \mathcal{B}/\mathcal{H},$$

where μ stands for the permeability. It expresses the fact that iron transmits the inductive effect better than air, or is more permeable. Magnetic induction is μ times the magnetizing force.

599. Magnetic Susceptibility. — The intensity of magnetization is the pole strength per unit area of the polar surface. Magnetic *susceptibility* is the ratio between the intensity of magnetization $\mathcal{I}$ and the strength of the field, or, in symbols,

$$\kappa = \mathcal{I}/\mathcal{H}.$$

The concept involved in permeability rather than the one in susceptibility is the modern one derived from Faraday.

600. Paramagnetic and Diamagnetic Substances. — A clear distinction between paramagnetic and diamagnetic substances may be drawn by means of their relative permeability as compared with that of air. Paramagnetic substances are those whose permeability is greater than unity; and since the permeability of air is practically unity, paramagnetic substances are more permeable than air. On the other hand, diamagnetic substances have a permeability less than unity, or they are less permeable than air.

Paramagnetic substances concentrate the magnetic flux and diamagnetic substances diffuse it. If iron be placed in a magnetic field, it will cause more lines of induction to pass through than through air; if bismuth be placed there instead of iron, fewer lines will pass through it than through the air previous to its introduction.

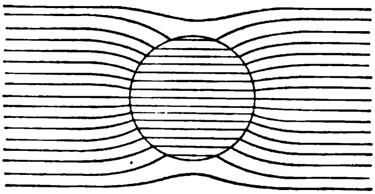


Fig. 374

If an iron sphere be placed in a uniform magnetic field, the effort of the lines of induction will be to run as far as possible through the sphere (Fig. 374). This action proceeds on the principle that the potential energy of a system always tends to as small a value as possible; for when the same magnetic flux passes through iron as through air, there is less energy per unit volume in the iron, for it requires a smaller magnetizing current to send the flux through.

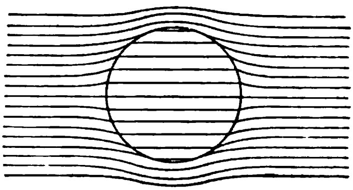


Fig. 375

If the sphere in Figure 375 is bismuth, the effort of the

lines of induction will be to avoid it because it is less permeable than air. For the same flux density, the energy per unit volume is greater in bismuth than in air.

When lines of magnetic induction pass from air into a paramagnetic substance, as in Figure 374, they are bent away from the normal to the surface in the substance; when they pass from air into a diamagnetic substance, they are bent toward the normal.

601. Movement of Paramagnetic and Diamagnetic Bodies in a Magnetic Field. — Faraday examined the magnetic behavior of a large number of bodies in the intense field between the pointed poles of a powerful electromagnet. A small bar of iron suspended between the poles (Fig. 376) turns in the axial direction *ab*, while a bar of bismuth sets its longer axis in the equatorial direction *cd* across the field. If the bismuth is in the form of a cube or a sphere, it is repelled to one side. Iron moves into the stronger parts of the field; bismuth into the weaker. They are examples of the two classes into which bodies are divided with respect to the action of magnetism on them.

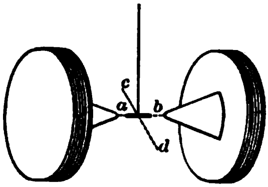


Fig 376

These movements may be satisfactorily explained by the relative permeability of the body and the medium in which it is suspended. Feebly magnetic bodies behave as if they were diamagnetic when surrounded by a more highly permeable medium. A small glass tube containing a weak solution of ferric chloride is paramagnetic in air; but when it is suspended in a denser solution of ferric chloride, it takes a cross position like a diamagnetic body. When any body assumes the equatorial position, the only inference which can justly be drawn from this behavior is that its permeability is less than that of the air or other medium surrounding it.

In general liquids are diamagnetic; liquid oxygen and solutions of salts of the paramagnetic metals are exceptions.

602. Curves of Magnetization. — The curves of Figure 377 show the relation between the magnetizing force and the magnetic induction for three samples of iron; *a* is the curve for mild steel, *b* for wrought iron, and *c* for cast iron. The magnetizing force $\mathcal{H}$ is plotted horizontally, and the induction $\mathcal{B}$ vertically; the resulting curves represent the successive stages of magnetization.

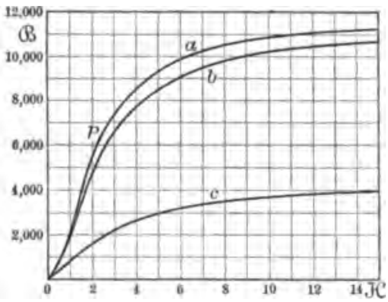


Fig. 377

If the ratio of $\mathcal{B}$ to $\mathcal{H}$ were constant, the curve of magnetization would be a straight line. For small magnetizing forces the curve is nearly straight; after this it bends sharply upward, and then becomes gradually flatter and flatter, so that for large values of the magnetizing force it is again a straight line. In the figure the scale for $\mathcal{B}$ is much smaller than for $\mathcal{H}$. If the same scale were used for the two, the curves for large values of $\mathcal{H}$ would be a straight line inclined to either axis at the angle of 45°. Under these conditions the iron is said to be saturated, and the permeability μ is a constant.

Since μ is the ratio of $\mathcal{B}$ to $\mathcal{H}$, it is evident that the permeability is a maximum for that point of the curve for which a straight line joining the origin and the point, as *P*, no longer cuts the curve, but is tangent to it. For smaller as well as for larger values of the coördinates the permeability is less.

603. Hysteresis. — If the magnetization of a ring of iron is carried through a complete cycle by increasing the magnetizing force by successive steps from zero to some definite

value, decreasing it from that value by small steps through zero to an equal value in the other direction, and then again reducing it to zero and completing the cycle, the curve connecting $\mathcal{B}$ and $\mathcal{H}$ will not be the same with decreasing values of $\mathcal{H}$ as with increasing ones (Fig. 378). The induction $\mathcal{B}$ lags behind the magnetizing force. Thus, when $\mathcal{H}$ is reduced to zero from its maximum positive value, $\mathcal{B}$ has the value *ob*, and $\mathcal{H}$ must be given a negative value equal to *oc* before the induction becomes zero. So when $\mathcal{H}$ returns from its maximum value in the other direction to zero, the induction decreases only to the value *oe*. This phenomenon of the lag of the induction behind the magnetizing force Ewing has called *magnetic hysteresis*. The result of plotting the corresponding values of $\mathcal{B}$ and $\mathcal{H}$ through a complete cycle is a curve inclosing an area, and *this area represents the heat lost per cubic centimeter in the iron in carrying it through a single cycle.*

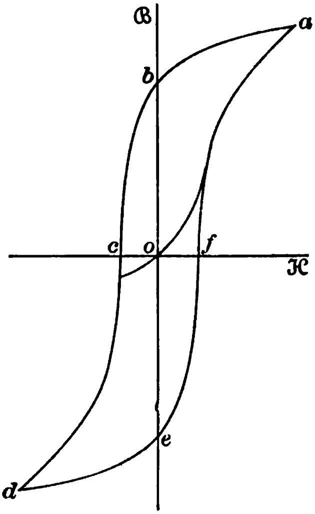


Fig. 378

604. Remanence and Coercive Force. — A cyclic magnetization curve serves among other things to give definiteness to the terms *remanence* or retentivity and *coercive force*. The residual value of $\mathcal{B}$ when $\mathcal{H}$ is reduced to zero is *ob* (Fig. 378). This value is the *remanence*. It depends on the quality of the iron, the limit to which the magnetization has been pushed, and whether the magnetic circuit of the iron is open or closed. The figure applies to a magnetic circuit consisting of a ring. The value of $\mathcal{H}$ required to reduce the residual induction to zero, namely, *oc*, is the

measure of the *coercive force.* Mechanical vibration due to external forces has the effect of diminishing residual magnetism, coercive force, and hysteresis. If the iron in thin plates be carried rapidly through successive cycles of magnetization by alternating currents, some vibration will be set up in the plates unless they are rigidly clamped together. Any vibration sustained by means of the current absorbs energy, and increases the area of the hysteresis curve.

VI. THE MAGNETIC CIRCUIT

605. Law of the Magnetic Circuit. — The idea of a magnetic circuit in a vague form is more primitive than that of an electric circuit, for it appears to go back to the mathematician Euler in 1761. Later Joule asserted that the resistance to magnetic induction is proportional to the length of a closed magnetic circuit. Faraday made the very apt comparison of an electromagnet with open magnetic circuit to a voltaic cell immersed in an electrolyte of poor conductivity. The low permeability of the air corresponds to the low conductivity of the electrolyte. Maxwell said, "In isotropic media the magnetic induction depends on the magnetic force in a manner which corresponds with that in which the electric current depends on the electromotive force."

The first definite expression of the law of the magnetic circuit in the form of an equation, like the equation expressing Ohm's law, was given by Rowland in 1873; he says expressly that it "is similar to the law of Ohm."

In 1883 Bosanquet introduced the term "magnetomotive force," corresponding to electromotive force in the electric circuit. We may then write

$$\textit{Magnetic flux} = \frac{\textit{Magnetomotive force}}{\textit{Magnetic reluctance}}.$$

Before attempting to write a more detailed equation for the magnetic circuit, it is necessary to introduce certain

general propositions which furnish an expression for the magnetomotive force.

606. Rotation of a Closed Circuit in a Magnetic Field. — Conceive a current of I *c.g.s.* electromagnetic units flowing through the half circle *abcd* (Fig. 379), and let there be a unit magnetic pole at the center P. Then the field produced at P by the current urges the pole in a direction normal to the plane of the ring (§ 580). The circuit is urged by an equal force in the opposite direction by Newton's third law.

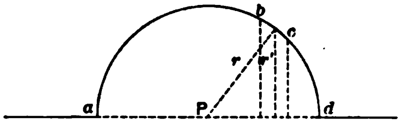


Fig. 379

Let *bc* be an element of the curve. Then by the reciprocal action between a magnet and a current, which has been experimentally demonstrated, the force f on the unit pole at P, due to the current I in 1 cm. of the circular conductor, is equal to I/r^2. Hence, the work done in rotating any short arc *bc* against this force through 360° about the axis *ad* is $f \times bc \times 2\,\pi r'$. But this expression is f times the area of that portion of the spherical surface generated by *bc* during the rotation. Therefore, the entire work done against the magnetic reaction between the current in the whole semicircumference and the unit pole at the center, for one revolution, is the product of f and the numerical value of the surface of the sphere whose radius is r, or

$$W = f \times 4\,\pi r^2 = \frac{I}{r^2} \times 4\,\pi r^2 = 4\,\pi I. \qquad (108)$$

Since $4\,\pi$ lines of force radiate from unit pole (§ 465), and all of these are cut by the semicircle during one rotation around the axis *ad*, it follows that *the work done is equal to the product of the whole number of lines cut by the conductor and the strength of the current flowing through it.*

607. The Electromotive Force Generated. — Assume the rotation described in the last article to take place in a period

of t seconds, that the resistance of the conductor between the points a and d is R, and that E is the applied potential difference between the same points. Then from the law of the conservation of energy, the whole work done is the sum of the energy spent in heating the conductor and the work done in rotating it in the magnetic field, or

$$EIt = I^2Rt + 4\,\pi I$$

as the energy equation.

Therefore $$E = IR + \frac{4\,\pi}{t},$$

and $$I = \frac{E - 4\,\pi/t}{R}. \tag{109}$$

This is an expression for the current in the form of Ohm's law. It shows that there is generated by the rotation an E. M. F. equal to $4\,\pi/t$. But this fraction is the rate at which the $4\,\pi$ lines of force from the unit pole are cut by the rotating conductor. *The E. M. F. generated by a conductor cutting across lines of magnetic force is, therefore, equal to the rate at which they are cut.*

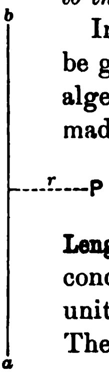


Fig. 380

In estimating the number of lines cut attention must be given to the direction in which they are cut, and the algebraic sum must be taken in all cases. Use will be made of this principle in the next chapter on induced electromotive forces.

608. Force at a Point due to a Current of Indefinite Length. — Let ab (Fig. 380) be a portion of the straight conductor carrying a current of strength I *c. g. s.* units, and let P be the point at a distance r from it. Then if a unit pole be at P and if the conductor be carried around it at the distance r, or the pole around the conductor at the same distance, all the lines of force from the pole will be cut once. The work done is $4\,\pi I$. Also, if the field produced by the current at the point

P is $\mathcal{H}$, the work done is the product of the magnetic force and the distance $2\pi r$, or $2\pi r\mathcal{H}$. Hence,

$$2\pi r\mathcal{H} = 4\pi I,$$

or,
$$\mathcal{H} = 2I/r.$$

If the current I is in amperes, the force in dynes at the point is

$$\mathcal{H} = 2I/10r.$$

609. Force within a Helix. — Let AB (Fig. 381) represent a section through the axis of a long helix, and let unit pole be at the point P. Let there be n turns of wire per centimeter parallel to the axis of the helix, each turn carrying a current of I *c. g. s.* units. Then if the unit pole be carried along the axis from P to P', a distance of one centimeter, each of the 4π lines from the pole will be cut by n turns of wire; the whole number of lines cut will be $4\pi n$, and the work done $4\pi nI$. Since the distance moved is one centimeter, the work is numerically equal to the force, or

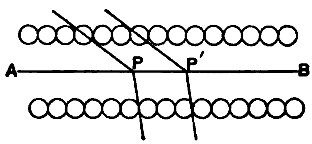


Fig. 381

$$\mathcal{H} = 4\pi nI. \tag{110}$$

If the current I is in amperes,

$$\mathcal{H} = 4\pi n\, I/10. \tag{111}$$

This is the value of the field of force at points distant from the ends of the helix. If the helix or solenoid forms a closed curve, so that there are no ends to the helix, the field along the magnetic axis is everywhere the same.

610. Magnetomotive Force. — The reader will recall that the electromotive force in a circuit is equal to the work required to carry unit quantity of electricity entirely around the circuit (§ 535). So magnetomotive force is equal to the work

done in carrying a unit pole once around the magnetic circuit. If L is the length of the solenoid, the work done is L times the strength of the field, or $4\pi nIL$, if the field is uniform. Now nL is the entire number of turns of wire in the solenoid. Denote this by N. Then the magnetomotive force is

$$\mathcal{F} = 4\pi NI/10 \tag{112}$$

if the current is expressed in amperes. The quantity NI is called the *ampere turns*. The magnetomotive force in a long solenoid is therefore 1.257 times the ampere turns.

Similarly, the difference of magnetic potential between two points of a solenoid is 1.257 times the number of ampere turns between the same points.

611. Reluctance.—The *magnetic reluctance* of a bar of iron may be calculated from its length, its sectional area, and its permeability, just as the electrical resistance of a conductor may be calculated from its length, its cross section, and its conductivity. Let the length of the bar be l centimeters, its section S square centimeters, and its permeability μ. Then its reluctance is

$$\mathcal{R} = l/\mu S. \tag{113}$$

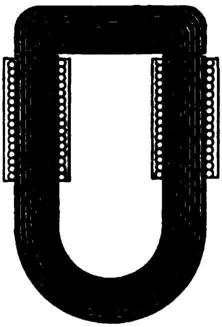


Fig. 382

For example, the reluctance of the magnetic circuit of the electromagnet of Figure 382 is made up of two parts, that of the core and that of the armature. Let the lengths, sections, and permeabilities be denoted by l_1 and l_2, S_1 and S_2, and μ_1 and μ_2 respectively. Then the reluctance of the whole circuit is

$$\mathcal{R} = \frac{l_1}{\mu_1 S_1} + \frac{l_2}{\mu_2 S_2}.$$

612. Law of the Magnetic Circuit Applied.—When the magnetic circuit is not closed, the lines of induction must be

forced across the air gap between the faces of the iron parts of the circuit. Suppose the armature removed a short distance l_3 from the poles (Fig. 383). Then the length of the circuit is thereby increased $2\,l_3$ cm., and additional reluctance is introduced equal to $2\,l_3/S_3$, where S_3 is the cross section of the air traversed by the induction. The permeability of the air is unity, and does not appear in the expression.

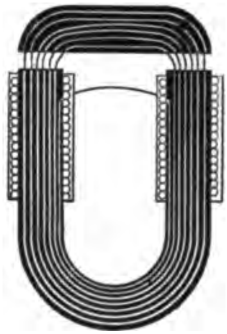


Fig. 383

We may therefore write for the flux of magnetic induction

$$\Phi = \frac{4\,\pi NI}{10\left\{\frac{l_1}{\mu_1 S_1} + \frac{l_2}{\mu_2 S_2} + \frac{2\,l_3}{S_3}\right\}}, \quad (114)$$

where I is expressed in amperes.

While this expression is simple in theory it is rendered difficult of application because μ, unlike conductivity, is not a constant, but is a function of the magnetization or induction in the iron. In applying the formula to any particular magnetic circuit it is necessary to know the curve of magnetization or the quality of iron used, and to ascertain from it or from tables the values of μ corresponding to the degree of saturation which it is desired to use. When this has been determined, the formula gives the number of ampere turns of excitation required. For open magnetic circuits an allowance must be made for leakage of lines of force through the air between parts of the magnet. This leakage requires excitation, but contributes nothing to the purpose for which the magnet is designed. The allowance for it must be estimated from experience with the particular form of magnet employed. The electromagnets of dynamos are designed by a process similar to this.

613. Superficial Magnetism by Electric Discharges. — Thin steel rods and sewing needles may be magnetized by passing

an electric discharge around them, or even across them at right angles to their length. It has long been known that hard steel is sometimes magnetized by lightning.

If a Leyden jar be discharged through a strip of tin foil across which lies a sewing needle, the needle will be magnetized by the discharge. Better results will be obtained by surrounding the needle with an open helix of rubber-covered wire and discharging through it. It was with simple means like these that Joseph Henry discovered the oscillatory character of the Leyden jar discharge.

Anomalous results have sometimes been observed in the relation of the poles to the direction of the discharge around the needles or rods, the poles being turned in the direction opposite to what the rule would lead one to expect. This result is due to the oscillatory discharge combined with the superficial character of the magnetism imparted. If small steel rods, magnetized by electric discharges, be examined by removing the external portions with acid, it will be found that the magnetized part is confined to a thin shell, the underlying parts remaining unmagnetized. If a second discharge succeeds the first in the opposite direction, it will reduce the external magnetism to zero if the magnetism of half the shell is reversed. Two shells of equal magnetic moment will then be superposed in opposite senses. If therefore the reverse discharge have more than half the magnetizing effect of the first, the resultant magnetism will be apparently "anomalous"; but it is accounted for by the direct and reverse discharges, and does not constitute an exception to the law of magnetization.

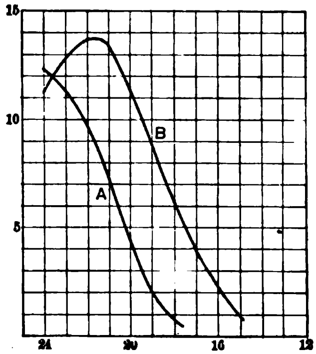


Fig. 384

Figure 384 contains the curves obtained from two glass-

hard steel rods, 6 cm. long and 1.8 mm. in diameter, magnetized by ten successive discharges of a small Leyden jar, all in the same direction. The relation of the two magnetizing coils was such that the first reverse oscillation was more powerful with *B* than with *A*. The data for these curves were obtained by removing successive portions of the outside with acid and measuring the magnetic moments after each removal. Moments are plotted as ordinates, and decreasing weights as abscissas. The moment of *B* at first increases to a maximum, and then decreases parallel to the *A* curve. *B* had a thin external shell magnetized in a sense opposite to that of the underlying portions. When this had all been removed, the magnetic moment was a maximum.

Problems

1. An iron bar 50 cm. long and 3 cm. in diameter was magnetized to 15,780 lines per square centimeter, when μ equaled 800. Find the reluctance and the total induction through the bar.

2. A ring of soft iron 20 cm. in diameter and 3 cm.^{2} sectional area is wound uniformly with a magnetizing helix. Find the number of ampere turns required to magnetize to 13,640 lines per square centimeter, with μ equal to 2200. What will be the total induction?

3. A straight wire carries a current of 10 amperes. Find the force in dynes on a pole of strength 20 at a distance of 5 cm. from the wire.

4. Find the strength of the magnetic field 8 cm. from the center of a coil of one turn in the line of its axis if the coil is 12 cm. in diameter and carries 0.5 ampere.

5. Two straight insulated conductors, indefinitely long, cross each other at right angles; one carries a current of 50 amperes and the other 100. Find the force on unit pole in their plane, at a distance of 6 cm. from the former and 8 cm. from the latter.

6. What current in amperes through a straight wire of indefinite length will produce at a distance of 10 cm. a field equal in dynes to the weight of 10 mg.

CHAPTER XX

ELECTROMAGNETIC INDUCTION

I. FARADAY'S EXPERIMENTS

614. Faraday's Discovery. — The discovery made by Oersted in 1819 led speedily to the discovery of magnetization by electric currents, and to the mechanical action between conductors conveying them. Faraday completed this correlated group of electromagnetic phenomena by discovering, in 1831, the laws of electromagnetic induction, that is, the laws of the production of induced currents by means of other currents or by magnets. These discoveries are of very great interest and importance, for all modern methods of generating large electromotive forces by dynamo machines, and all induction coils and alternate current transformers, are based on the principles of electromagnetic induction.

Induced electromotive forces and currents are those produced by the action of magnets and other currents. Strictly only electromotive forces are induced; induced currents flow as a consequence when the circuit in which the electromotive force is generated is closed. But electromotive forces or potential differences may be induced just the same when the circuit is open.

615. Induction by a Magnet. — Assume a coil of many turns of insulated wire connected in circuit with a sensitive galvanometer (Fig. 385), and that the north pole of a bar magnet is quickly thrust into the coil. The galvanometer will indicate a transient current, which flows only during

the motion of the magnet. If the magnet be suddenly withdrawn from the coil, a transient current will flow in the reverse direction. If the south pole of the magnet be thrust into the coil and then withdrawn, the currents in both cases will be the reverse of those with the north pole.

Fig. 385

The momentary electromotive forces generated in the helix are known as *induced electromotive forces*, and the currents as *induced currents*. The magnet carries with it into the coil its lines of induction; and when the relative position of a magnet and a coil are so altered that a variation is produced in the induction linked with the coil, an induced electromotive force is generated in it.

If a coil of fine wire be wound around the armature of a permanent magnet (Fig. 386), when the armature is in contact with the poles the flux of induction through the coil is a maximum. When it is pulled away, the magnetic flux through the armature and the coil decreases rapidly, and an E. M. F. is generated. This experiment illustrates Faraday's method of producing electric currents by the aid of magnetism.

Fig. 386

When the armature is in contact with the magnet, and one or more fine iron wires are dropped so as to join the poles, the galvanometer, if sufficiently sensitive, will show an induced E. M. F. The iron wires divert magnetic flux from the armature by forming a divided magnetic circuit, and the sudden decrease of

induction through the armature generates an E. M. F. in the coil surrounding it.

616. Direction and Value of an Induced Electromotive Force. — The direction of the induced E. M. F. Faraday determined by experiment, but it can be deduced from considerations with which we are already familiar.

Suppose a magnet *NS* thrust into a helix (Fig. 387). Since an E. M. F. is generated and a current circulates through the coil, the energy of the current must be derived from the work done in moving the magnet. There must therefore be resistance opposing the movement; this resistance is magnetic. It is due to the helix considered as a magnetic shell, and the current must flow around it in a direction to make a north pole of the side which the north pole of the bar magnet enters. Its direction is therefore against watch hands, as indicated by the arrows. If the observer looks along the positive direction of the lines of induction, a current flowing clockwise is said to be *direct;* if counterclockwise, it is *indirect.* We have therefore the following law relating to the direction of the induced electromotive force:

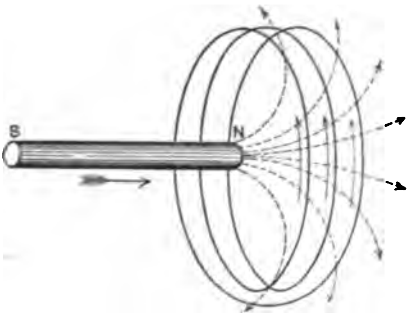


Fig. 387

An increase in the number of lines of induction threading through a helix produces an indirect electromotive force; a decrease in the number of lines produces a direct electromotive force.

The numerical value of an induced electromotive force in *c.g.s.* units may be expressed as follows:

The electromotive force induced is equal to the rate of change in the number of lines of induction threading through the circuit.

If $d\Phi$ is the change in the magnetic flux through the circuit taking place in the short time dt, the induced E. M. F. is

$$E = -d\Phi/dt.$$

The minus sign indicates that a direct electromotive force corresponds to a decrease in the flux of induction.

617. Induction by Currents. — Since a current through a solenoid produces a magnetic field equivalent to that of a magnet, the same induction effects will be produced by inserting a helix conveying a current into the long coil (Fig. 388) as by thrusting in the equivalent magnet. Let the circuit P include a battery and a key, and the circuit S a galvanometer. The former is called the *primary* and the latter the *secondary*.

Fig. 388

If the current through P is kept constant while the coil is moved about, when P approaches S, an E. M. F. is generated in S tending to send a current in the opposite direction to that around P; while if P is moved away from S, the E. M. F. induced in S is in the direction of the current around P. These electromotive forces in S act only so long as P is moving. If P is kept fixed while S is moved, the results are the same.

Next, let P be in a fixed position near S with the key open. Then on closing the key in P the galvanometer

needle will be deflected. This deflection is not a permanent one, but the needle oscillates and finally returns to its initial position of rest, indicating the passage of a sudden discharge through the galvanometer. The direction of this momentary current is opposite to that through P. On opening the key another similar momentary current passes through S, but in the same direction as through P. Thus the starting or stopping of a current in P is accompanied by the induction of another current in a neighboring circuit S.

The sudden increase of the current in P produces an opposite current in S, and the sudden decrease of the current in P produces a current through S in the same direction as through P.

If when P remains inside of S, or coaxial with it, a bar of soft iron is placed within it, there is an increase of magnetic flux through both P and S, and the E. M. F. generated in S is in the same direction as that produced by closing the key in P, moving P toward S, or increasing the current through P. The withdrawal of the iron produces the opposite effects to its insertion in the coil.

The law of the direction and magnitude of the E. M. F. generated inductively by another current is the same as that given in the last article. When the magnetic flux changes, an E. M. F. is produced equal to the rate of change in the magnetic flux passing through the circuit. The positive direction of the E. M. F. and of the flux through the circuit are related to each other as are the rotation and the translation of a right-handed screw.

618. Faraday's Ring. — Faraday wound upon an iron ring two coils of wire P and S (Fig. 389). When a battery and a key were included in the circuit P and a galvanometer in S, whenever the circuit of P was closed or opened a momentary current was produced in the closed circuit S. In this experiment the iron is the medium through which the induction between P and S takes place. The current

through P magnetizes the iron ring as a closed magnetic circuit. The starting of the current in the circuit P sends magnetic lines through S and produces in it an inverse current; the stopping of the primary current withdraws lines and produces a direct current through the secondary. A larger deflection of the galvanometer will be produced by the first closing of the primary circuit than by opening it, or by closing it a second time unless the current is reversed. The reason is that the ring forms a closed magnetic circuit, and its retentivity or remanence is so great that only a small part of the lines of induction drop out when the magnetizing current ceases to flow. But if the current through the primary be reversed, all the lines will be taken out and will be put in again the other way round. Hence, a large induction will take place in S. A closed magnetic circuit is not well adapted, therefore, to produce induction effects by merely opening and closing the primary circuit.

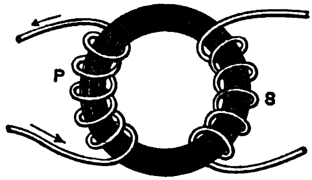


Fig. 389

The relation between P and S is a mutual one. If S is made the primary, induced electromotive force will be generated in P as the secondary. The Faraday ring with its two coils is the type of the modern transformer for alternating currents.

619. An Inductive System a Conservative System. — It is instructive to consider a system of two circuits, or of a circuit and a magnet, as a conservative system. The action between the parts of the system always tends to maintain unchanged the number of lines of induction threading through the circuits. Thus, in Figure 387 the approach of the magnet to the coil increases the magnetic flux through the coil, and the induced current is in a direction to send a counter flux through it so as to keep the flux linked with the coil a constant. In Figure 389 the primary current pro-

duces magnetic flux in the ring, and the current induced in the secondary produces magnetic flux in the other direction around the ring; that is, the induced current opposes any change in the flux. After the primary current has produced a steady magnetic flux through the iron, the opening of the primary circuit induces a secondary current in the same direction around the ring as the primary, and this tends to maintain the flux of induction in the ring unchanged.

The same principle may be applied to two coils without iron. There is no exception to the law that induced currents are always in a direction to conserve the status of the magnetic flux through the circuit in which the induction takes place. This law means that the magnetic flux through a circuit does not change abruptly—a property of magnetic induction analogous to inertia in matter.

620. Lenz's Law. — When induced currents are produced by the motion of a conductor in a magnetic field, the circuit is acted on by a mechanical force. Lenz's law is that the direction of this force is such that the force opposes the motion which induces the current. Lenz's law is a particular case of the principle of the conservation of magnetic flux. Every action on an electromagnetic system which involves a transformation of energy sets up reactions tending to preserve unchanged the state of the system.

An example of Lenz's law is afforded by a coil revolving in a magnetic field. The mechanical action of the field on the current induced in the coil produces a couple tending to stop the rotation. The oscillations of the coil of a D'Arsonval galvanometer subside quickly when the coil is short-circuited. The galvanometer is then a magneto-electric machine, and the currents induced in its closed coil bring it to rest.

Broadly stated, Lenz's law is as follows:

The direction of an induced current is always such that it produces a magnetic field opposing the motion or change which induces the current.

621. Arago's Rotations. — When a magnet is suspended horizontally over a copper disk and the disk is rotated, induced currents are produced in it. These give rise to a force opposing the rotation. Since the force between the magnet and the disk is a mutual one, a couple acts on the magnet and turns it, if it is free to move, in the same direction as the disk. Or if the magnet is spun around a vertical axis and the disk is movable, it is dragged after the magnet. These motions are called Arago's *rotations;* they were discovered by Arago, but were first explained by Faraday. Induced currents flow in closed circuits through the disk, and the action between them and the magnet tends to stop the disk; or if the magnet oscillates, the induced currents damp its motion. Thus in Figure 390, if the needle *ab* oscillates over the disk, as it moves in the direction of the arrows, a current is induced on the *M* side which repels the needle, and one on the *N* side attracting it; or the current under it flows from the center toward the circumference if *a* is an N-seeking pole.

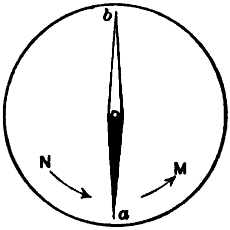


Fig. 390

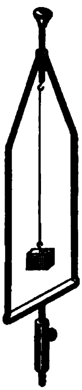


Fig. 391

622. Other Examples of Lenz's Law. — Let a copper cube or cylinder be suspended between the pointed poles of a powerful electromagnet (Fig. 391). The cube may be set rotating by twisting the thread and releasing it. When the electromagnet is excited the cube is instantly brought to rest; it begins to spin as soon as the magnetizing current is cut off, and is again arrested when the circuit is closed. This resistance to motion in a magnetic field is sometimes called magnetic friction.

In another experiment a disk of copper is made to rotate rapidly between the poles of an electromagnet (Fig. 392). When the magnet is excited, the disk appears to meet with a sudden resistance. Foucault found that if it is forced to rotate, it is heated by the induced currents flowing in it. These induced currents in masses of metal are

often called Foucault currents. There is a pair of eddy currents in the part of the disk passing the poles; and these currents, as in Arago's rotation, hold the disk back.

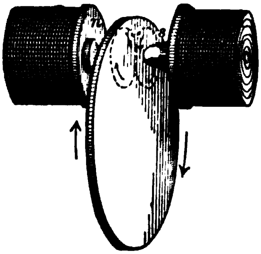


Fig. 392

The drag due to eddy currents is proportional to the speed and to the square of the magnetic field; for the force is proportional to the field and the current, and the current is proportional to the field and the speed. When the field is constant, the force is therefore proportional to the speed of rotation.

The principle is employed to produce damping in rotatory meters. A copper disk, attached to the shaft to which is connected a dial train, rotates between the poles of fixed magnets. The drag on the copper disk keeps the speed proportional to the torque.

II. SELF-INDUCTION

623. Self-induction. — Joseph Henry discovered that a current through a helix with parallel turns acts inductively on its own circuit, producing what has long been known as the *extra current*, and a bright spark across the gap when the circuit is opened. The effects are not very marked unless the helix contains an iron core.

Even a single circuit is a conservative system as regards the magnetic flux through it. When the current magnetizes the core, the effect is the same as if a magnet had been plunged into the helix; that is, the induced E. M. F. is a counter E. M. F. tending to prevent the flux of magnetic induction through the circuit. The result is that the current in such a circuit does not reach its maximum value abruptly, but only after a short interval depending on the value of the *coefficient of self-induction*, or simply the *inductance*. On the other hand, when the circuit is opened the induced E. M. F. is direct and tends to prolong the current, or to resist the diminution in the magnetic flux.

The unit of inductance, the *henry*, is the inductance in the circuit when the E. M. F. induced in this circuit is one volt,

while the inducing current varies at the rate of one ampere per second.

624. Growth of Current in Inductive Circuits. — When a constant electromotive force is impressed on a circuit having self-inductance, the current does not attain its permanent value instantly. During the variable stage its value is not given by the simple application of Ohm's law; the inductance is another property of the circuit, in addition to its resistance, which determines the instantaneous value of the current. When the circuit is closed the self-induced electromotive force opposes the applied electromotive force and retards the growth of the current to its full value. When the circuit is opened, the self-induced electromotive force prolongs the current, or acts to retard its decrease to zero.

Figure 393 is reproduced from a graph made by the current itself by means of an instrument called an oscillograph.* The oscillograph is in principle like a D'Arsonval galvanometer but with an extremely short period of oscillation, so that the moving system is able to follow all the changes of the current. A reflected beam of sunlight is received on a falling photographic plate and thus records the motion of the mirror. The vertical line is the trace of the beam immediately after closing a noninductive circuit; the other line is the trace obtained with a parallel inductive circuit. The opposing electromotive force of self-induction is greatest at the instant the circuit is closed; as it dies away, the effective electromotive force increases, and the current rises to the value given by Ohm's law.

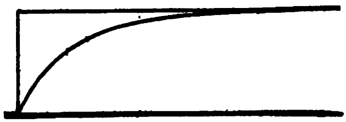


Fig. 393

625. Energy stored in a Magnetic Field. — The inductance is the property of a circuit by virtue of which the passage of a current is accompanied by the absorption of energy in the form of a magnetic field. If no other work is done, part of the energy flowing from the source is converted into heat,

*My thanks are due to Professor Benjamin F. Thomas for his beautiful oscillograms from which this illustration and a few subsequent ones were made.
—H.S.C.

and the rest is stored in the ether as the potential energy of the field. This storage of energy goes on while the current is rising from nothing to its steady value. The work represented by this energy is done by the current against the electromotive force of self-induction.

The storage of energy may be strikingly illustrated as follows: *M* (Fig. 394) is a large electromagnet, *B* a storage battery, *L* an incandescent lamp of a normal voltage equal to that of the battery, and *K* a circuit breaker. The circuit is divided between the electromagnet and the lamp; and since the former is of low resistance, when the current reaches its steady state most of it will go through the coils of the magnet, leaving the lamp at only a red glow. At the instant when the circuit is closed, the self-induction of the magnet acts against the current, like a large resistance, and sends most of it around through the lamp. It accordingly lights up at first, but quickly grows dim as the current rises to its steady value through *M*.

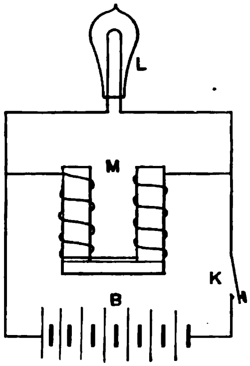


Fig. 394

On breaking the circuit and cutting off the battery, the lamp flashes up brightly. The lamp and the electromagnet are then together on a closed circuit. The energy stored in the magnetic field, as a strain in he ether about the magnet, is converted into electric energy, and a reverse current through the lamp lights it momentarily.

III. THE INDUCTION COIL

626. The Induction Coil. — An *induction coil* is commonly employed to obtain transient flashes of high electromotive force in rapid succession. In modern terms it is a step-up transformer with open magnetic circuit. About an iron core, consisting of a bundle of fine iron wires to avoid induced or eddy currents in the metal of the core, is wound a primary coil of comparatively few turns of stout wire; outside of this, and carefully insulated from it, is the secondary of a very large number of turns of fine wire. In Spottis-

wood's great coil, which gave $42\frac{1}{2}$-inch sparks, the secondary contained 280 miles of wire wound in 340,000 turns.

The primary must be provided with a circuit breaker (Fig. 395), if the coil is to be used with direct interrupted currents. It is commonly made automatic by a vibrating device actuated by the core and similar to that of a vibrating electric bell.

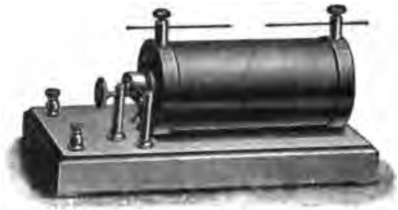


Fig. 395

In large coils the secondary is wound in flat spirals, and these are slipped on over the primary and separated from one another by insulating rings. The difference of potential between adjacent turns of wire is then not so large as when the entire coil is wound in layers from end to end, and it is easier to maintain the insulation. The ratio of the transformation of the electromotive force is nearly the same as the ratio between the number of turns of wire on the primary and the secondary.

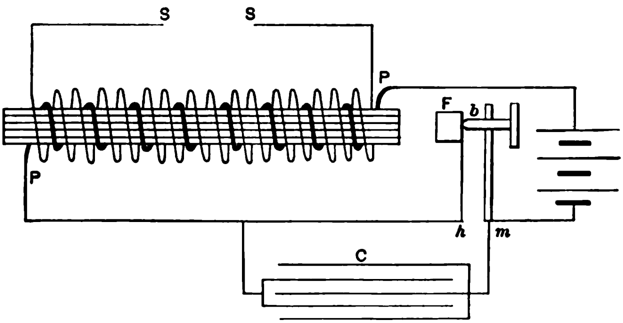


Fig. 396

627. Action of the Coil. — The essential parts of an induction coil are shown in Figure 396. The current from the battery passes through the heavy primary wire *PP*, thence through the spring *h*, which carries the soft iron block *F*,

then across to the platinum-tipped screw *b*, and so back to the negative pole of the battery. The attraction of *F* by the magnetized core breaks the primary circuit at *b*; the core is then demagnetized, and the release of *F* again closes the circuit. Electromotive forces are induced in the secondary coil *SS*, both at the make and break of the primary.

The self-induction of the primary has an important bearing on the action of the coil. At the instant the circuit is closed, the counter E. M. F. opposes the battery current, and prolongs the time of reaching its greatest strength. Consequently the E. M. F. of the secondary is diminished by the self-induction of the primary. The E.M.F. of self-induction at the break of the primary is direct, and this added to the E. M. F. of the battery produces a spark at the break points.

628. The Condenser. — The addition of a condenser increases the E. M. F. of the secondary coil in two ways: *First.* It gives such an increase of capacity to the primary coil that at the moment of breaking the circuit the potential difference between the contact points at the break does not rise high enough to cause a spark across the opening. The interruption of the primary is therefore more abrupt, and the E. M. F. of the secondary is increased. *Second.* After the break, the condenser *C*, which has been charged by the E. M. F. of self-induction, discharges back through the primary coil and the battery. The condenser causes an electric recoil in the current, and returns the stored charge as a current in the reverse direction through the primary, thus demagnetizing the core, increasing the rate of change of magnetic flux, and increasing the induced E. M. F. in the secondary. The condenser momentarily stores the energy represented by the spark which occurs without the condenser, and then returns it to the primary and by mutual inductance to the secondary, as indicated by the longer spark or the greater current. When the secondary terminals are separated, the discharge is all in one direction and occurs when

the primary current is interrupted. With a suitable condenser the conditions are those described by the word *resonance*. The current through the primary is rendered oscillatory by means of the condenser.

629. The Tesla Induction Coil. — Tesla's device to produce electric discharges of very high frequency employs inductive processes in two stages; in the second stage the oscillatory discharge of a Leyden jar (§ 521) serves as the interrupter. The terminals of the secondary of an induction coil *C* (Fig. 397) are connected to the inner and the outer coating respectively of a Leyden jar *J*. The discharge circuit of the jar is through the primary winding of the Tesla coil and the discharge balls S_1. The primary *A* of the Tesla coil consists of a few turns of heavy wire without a magnetic core. The secondary *B* has a much larger number of turns and it is separated from the primary by air or oil insulation. The oscillations of the Leyden jar discharges at S_1 may have a frequency of several millions per second. The discharges at S_2 from the secondary of the Tesla coil are not only of very high frequency, but they are of such high E. M. F. that they produce auroral displays of astonishing brilliancy and marked induction effects at a distance of several feet.

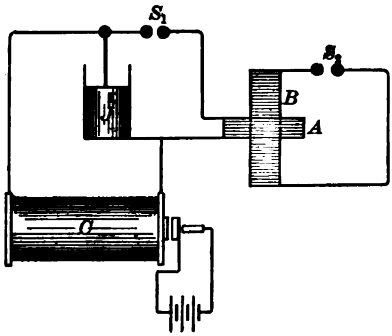


Fig. 397

IV. THE TELEPHONE

630. The Telephone. — The telephone was invented by Graham Bell and Elisha Gray in 1876. It consists of a permanent magnet *O* (Fig. 398), one end of which is surrounded

by a coil of many turns of fine insulated copper wire with its ends connected to the binding posts *t* and *t*. At a distance of a millimeter or less from the pole of the magnet on which the coil is wound is an elastic iron diaphragm or disk *a*; it is kept in place by the conical mouthpiece *d*.

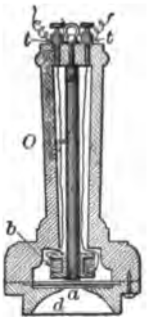


Fig. 398

Sounds may be transmitted to a distance by the use of two telephones, using one as a transmitter and the other as a receiver. The two wire coils are connected in series by the line wire between the two stations.

The diaphragm is magnetized by induction, and this induced magnetism reacts on the permanent magnet *O*, the amount of the reaction depending on the distance between the diaphragm and the surface of the pole. When words are spoken into the mouthpiece, the vibrations of the air cause the diaphragm to vibrate in unison; and the to-and-fro motion of the disk causes the induction between it and the magnet to vary in unison with the vibrations of the air. The variation in the number of lines of induction through the coil *b* produces a corresponding series of induced currents through the circuit in which the coil is connected.

When these varying induced currents, alternating in direction, traverse the coil of the receiving instrument, the magnetic field due to the coil, combined with that due to the magnet, alter intermittently the force between the magnet pole and the disk; the disk is therefore set in vibration in such a way as to reproduce the vibrations in the diaphragm of the transmitting instrument. The vibrations of the diaphragm of the receiving instrument are communicated to the air in contact with it, and thus the sounds made near the transmitter are reproduced by the diaphragm of the receiver.

631. The Microphone. — The currents produced by induction in a telephone as a transmitter are very feeble and serve

to transmit the voice over short distances only. The variable resistance transmitter, discovered by Berliner in 1877 and independently by Hughes in 1878, vastly improved the efficiency of transmission by causing the vibrations of the diaphragm to vary the resistance of a battery circuit by means of one or more contacts, the resistance of which varies with the pressure. The first recorded and definite transmission of speech was in fact accomplished by the variable resistance transmitter of Elisha Gray as early as 1876. It employed the simple device of a fine platinum wire soldered to the center of a diaphragm and dipping a little way into an acid solution. The motion of the diaphragm varied the depth of immersion of the wire and so the resistance of transmission by the electrolyte.

One form of the *microphone*, invented by the English electrician Hughes, consists of a rod of gas carbon *A* (Fig. 399), pointed at both ends and resting lightly in conical hollows made in blocks *CC* of the same material attached to a sounding board. These blocks are placed in circuit with a battery and a telephone receiver by means of the wires *X* and *Y*. The least disturbance of the sounding board, such as that caused by the ticking of a watch lying on the base *D*, disturbs the rod and the pressure with which it rests against the upper carbon block. Since the resistance of a carbon contact varies with the pressure, the disturbance of *A* causes variations in the resistance of the circuit, and therefore a continuously varying current. The receiver responds to the varying current; and so great is the sensitiveness of the instrument that a fly may be heard walking on the sounding board.

Fig. 399

632. The Solid Back Transmitter.— Instead of the loose carbon contact of the early microphone, carbon in granules between carbon plates is now extensively used. The form of transmitter employed for long distance transmission is the "solid back" transmitter (Fig. 400). The figure shows only the essential parts in section, with minor details omitted. *M* is the mouthpiece, and *F* and *C* the front and back parts of the metal case. The aluminum diaphragm *D* is held around its circumference by a soft rubber ring. The block *W* has a recess in front to receive the carbon electrodes *A* and *B*. Between them are the carbon granules. The block *E* is attached to the diaphragm and is insulated from *W* except through the carbon granules. The transmitter is placed in circuit by wires connected to *W* and *E*.

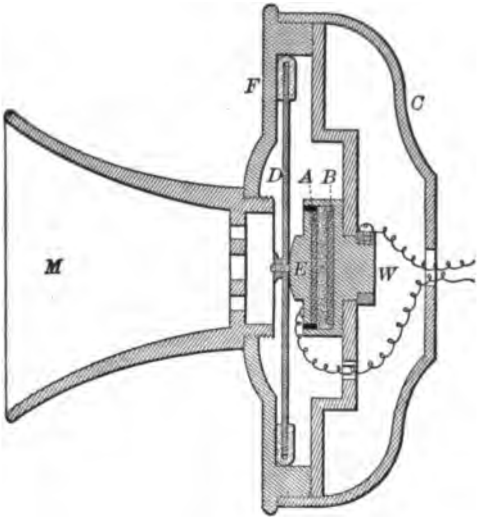


Fig. 400

Provision is made for an elastic motion of the diaphragm and block *E*. Sound waves striking the diaphragm cause a varying pressure between the plates and the carbon granules; this varying pressure varies the resistance offered by the granules and so varies the current. The transmitter is in circuit in the line with the primary of a small induction coil, the secondary being in a local circuit with the telephone receiver. The induced currents in the secondary have all the modifications of the primary current; and when they pass through a receiver, it responds and reproduces sound waves similar to those which disturb the diaphragm of the transmitter.

CHAPTER XXI

DYNAMO-ELECTRIC MACHINES

I. DIRECT CURRENT MACHINES

633. Ideal Simple Dynamo. — A dynamo is a machine for converting mechanical energy into the energy of an electric current. It is a generator of electromotive force, and is based on the principles of electromagnetic induction discovered by Faraday. It consists of a system of conductors, called an *armature*, revolving in a magnetic field in such a manner as to vary continuously the magnetic flux through them.

Suppose a single loop of wire revolving in a uniform magnetic field between the poles N S of a magnet (Fig. 401) around a horizontal axis in the direction of the arrow. The loop of wire in the position shown in the figure incloses the maximum magnetic flux Φ. When it has revolved through an angle θ, the

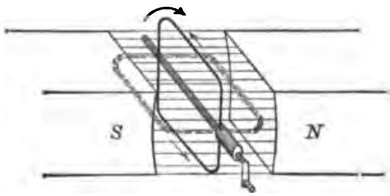


Fig. 401

flux through it will be reduced to $\Phi \cos \theta$; for the projection of the loop on a plane perpendicular to the field varies as the cosine of the angle of displacement from that plane. After a quarter turn the loop does not inclose any lines of induction; as it revolves further they thread through in the opposite direction, and this is equivalent to a continued diminution of the magnetic flux through the loop. During

the second half of the revolution the opposite changes take place; when the loop has revolved through 360° it returns to its initial relation to the magnetic field.

The magnetic flux through the loop varies therefore as the cosine of the angle of displacement θ. During the first half revolution a current flows around the loop in the direction of the arrows; during the second half it is in the reverse direction. Thus the E. M. F. changes sign twice every revolution. Such a loop, or coil composed of a number of parallel turns, generates an alternating electromotive force.

634. Law of the Electromotive Force. — The induced electromotive force is equal to the rate of change of the magnetic flux through the circuit (§ 616). But the flux inclosed by the loop varies as the cosine of the angle defining the position of the loop; and when the flux is a maximum, its rate of change is a minimum and conversely. Hence when θ is zero or 180°, the E. M. F. generated is zero; while for the positions 90° and 270° the E. M. F. is a maximum. The trigonometrical function that is related in this way to the cosine is the sine. Hence the law of variation of the electromotive force generated by the revolution of the loop in a uniform magnetic field is the same as the variation in the value of the sine of the angle of position.

If, therefore, we plot uniform distances along a straight line to represent equal increments of θ, and erect perpendicu-

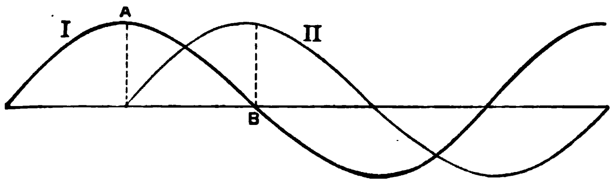


Fig. 402

lars to denote the values of the corresponding sines of θ, the curve connecting the extremities of these ordinates will be a *sine curve*. In Figure 402 the heavy line I is the cosine curve,

representing the law of change in the magnetic flux through the loop; the light line II is the sine curve, whose ordinates denote the rate of change of the flux, or the E. M. F. Their maximum values differ in phase by 90°, or a quarter of a period. When the magnetic flux decreases through its zero value at *B*, its rate of change is greatest and the electromotive force is a maximum and positive.

635. The Commutator. — The current in a single loop, or coil composed of several parallel turns of wire, has alternately equal values in opposite directions through the loop. To make it unidirectional in the external circuit a two-part *commutator* must be used. The two halves of a split tube (Fig. 403), insulated from each other and mounted on the shaft of the armature, are connected to the two terminals of the rotating coil. The brushes leading to the external circuit are so placed that they exchange contact with the two commutator segments in passing through the positions where the current changes its direction in the coil. The pulses are then all in one direction in the external circuit. Alternate loops of the sine curve are thus reversed, so that all of them lie on the same side of the zero line. One of the brushes is then always positive and the other negative.

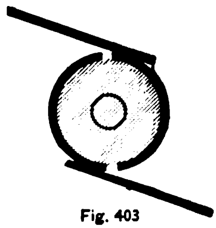


Fig. 403

When the armature consists of two loops or coils placed at right angles to each other, the commutator has four segments, and the frequency of the pulses in the external circuit is doubled.

636. The Gramme Ring. — The use of a commutator with more than two segments is conveniently illustrated in connection with the *Gramme Ring* (Fig. 404). The armature has a core made either of iron wire or of thin disks at right angles to the axis of rotation. The iron is divided for the purpose of preventing eddy currents in it, which heat the

machine and waste energy. A number of coils are wound in one direction and are all joined in series in a closed circuit. Each junction between coils is joined to a commutator bar. When a coil is in the highest position in the figure, the maximum flux passes through it; as the ring rotates, the flux through the coil decreases, and after a quarter of a revolution there is no flux through it.

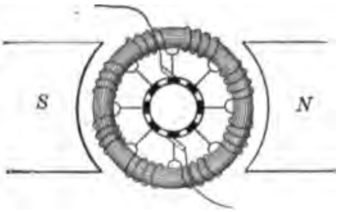


Fig. 404

The current through each coil reverses twice during each revolution, exactly as in the case of a single loop, when the magnetic field has only two poles. No current flows entirely around the armature, because the E. M. F. generated in one coil at any instant is exactly counterbalanced by the E. M. F. generated in the coil opposite.

When the external circuit connecting the brushes is closed a current flows in opposite directions through the two halves of the armature. The current through the armature has, in other words, two paths, and one brush is constantly positive and the other negative.

By increasing the number of coils, the potential difference between the brushes never drops to zero, as it does with a single coil twice every revolution, but it is kept nearly constant and at the highest value given by half the coils in series. The brushes must bear on the commutator near the part of the field where the E. M. F. in any coil passes through zero value and reverses.

637. The Field Magnet. — In direct current machines the magnetic field in which the armature rotates is produced by a large electromagnet excited by a current from the armature. After this field magnet has once been magnetized, the residual magnetism is sufficient to start the induced current, which is carried wholly or in part around the field magnet coils.

When the entire current passes through the coils of the field magnet, the dynamo is said to be *series-wound*. When the field magnet is excited by coils of many turns of wire connected as a shunt to the external circuit, the dynamo is said to be *shunt-wound*. The two methods of connecting the field magnet with the armature are shown in Figure 405. The

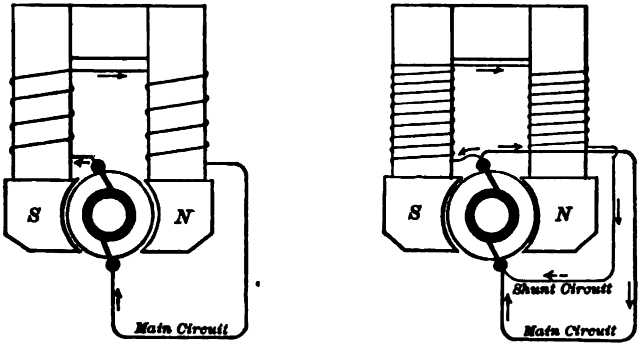


Fig. 405

series-wound machine is adapted to furnish a constant current with varying potential only; the shunt-wound machine is designed for a constant potential with varying current. In the latter when the current changes as a result of a change in the external resistance, the excitation of the field magnet remains nearly the same and the E. M. F. generated is nearly constant.

A *compound-wound* dynamo consists of a combination of shunt and series coils on the field magnet. It is designed to maintain the potential difference between the brushes more nearly constant than is possible with a shunt machine. The larger the current through the armature, the larger the loss of potential in the armature itself, the smaller the potential difference between the brushes, and the smaller the excitation of the field magnet. By carrying the whole current around the field magnet in a series coil of a few turns, the increased excitation thus produced compensates for the loss of potential

in the armature and maintains a constant potential difference between the brushes. If the armature were without resistance, compounding would not be necessary except for the demagnetizing effect of the armature considered as an electromagnet.

638. The Drum Armature.— The modern *drum armature* for direct currents is an evolution from a single coil shuttle armature. It is shown at *A* (Fig. 406), which represents the parts of a four-pole machine.

In the drum armature the iron core, which is made up of laminated disks, contains a series of grooves, parallel to the

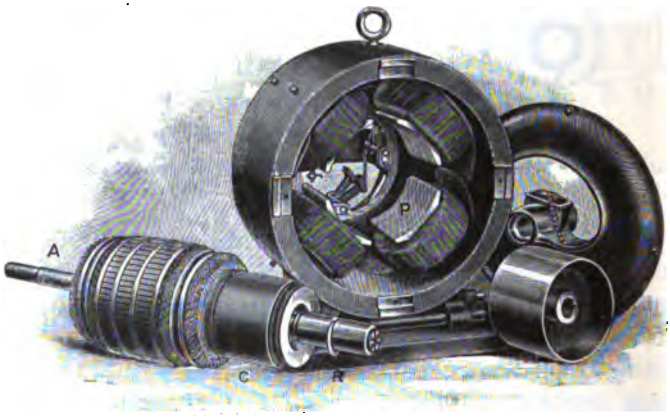


Fig. 406

shaft, and coils wound in them at equal angular distances around the circumference. These sections of the armature may all be joined in series, and the junctions between them are then connected with the commutator bars *C*, as in the Gramme ring. The machine of Figure 406 has four poles, four parallel circuits through the armature, and four brushes. Two of the brushes are positive and two negative; the two positives are connected in parallel as the positive terminal, and the two negatives as the negative terminal. With such

a winding, divided into numerous sections with corresponding commutator bars, the current in the external circuit is almost perfectly continuous and free from small fluctuations.

639. Reactions in the Field of a Generator and a Motor. — An electric motor for direct currents is constructed in the same manner as a generator. The study of a magnetic field through which a current is passing throws much light on the interactions between the field and the armature current. Figure 407 is a reproduction of a photograph of the field between unlike poles distorted by a current through a loop of wire which came up through one hole and went down through the other. The lines of induction, outlined by means of fine iron filings, are so distorted that some of them thread through the loop.

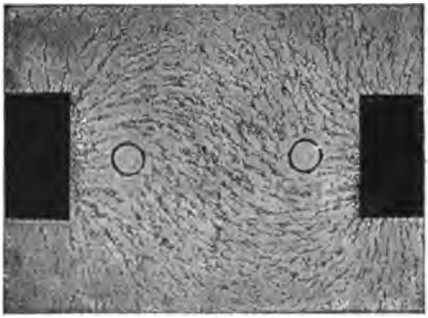


Fig. 407

If we conceive this loop to rotate clockwise around an axis perpendicular to the plane of the paper, it is clear that mechanical force must be applied to keep up the motion, because the tension along the lines of induction holds the loop back. The armature then turns against the magnetic forces or torque of the field acting on it, and operates as a generator, and the field of the machine is distorted in the direction of the rotation.

If, on the contrary, this loop of wire be conceived to rotate as an armature in the direction of the magnetic stress on it, then with the current in the same direction the field remains the same, and the armature reverses its direction of rotation.

When the machine is used as a generator mechanical power is converted into electrical energy, because the rotation of

the armature is maintained against the internal magnetic reactions in the field. Work is then done on the machine as a generator. On the other hand, when it is used as a motor, electrical energy is converted into mechanical work, because the rotation takes place in the direction of the magnetic effort or torque between the field and the armature. Work is then done by the machine as a motor.

640. Direction of Rotation as a Motor. — A series machine when used as a motor runs in the opposite direction to its motion as a generator. It will rotate in the same direction whether the current goes through it in one direction or the other, since it is reversed through the armature when it is reversed through the field.

A shunt machine runs in the same direction as a motor and as a generator. If in the shunt machine of Figure 405 the current from an external source enters the lower brush, as in the figure, its direction through the armature remains unchanged; through the field coils, however, it goes in the opposite direction to the arrows, for the armature and the field are now in parallel with reference to the external circuit or source of current. When used as a generator, the external circuit and the field are in parallel with respect to the armature as the source of the electric pressure. The field is therefore reversed. But as a motor the machine runs *with* the magnetic torque, and as a generator *against* it; so that running with the torque when the field only is reversed is the same as running against it before the field is reversed. For this reason the shunt-wound machine runs in the same direction whether used as a motor or a generator.

641. Counter Electromotive Force in a Motor. — The armature of an electric motor revolves in a magnetic field and generates an E. M. F. A little consideration will show that this E. M. F. must be an opposing one tending to reduce the current through it. In Figure 408 a generator and a motor are shown connected together. The direction of rotation of the two machines is the same. The direction of the electromotive forces generated in the two armatures is shown by the arrows. They are toward the lower brush in both, because both armatures revolve in the same direction in similar fields. But in the generator the current is in

the same direction as the E. M. F. generated in the armature, while in the motor it is against this generated E. M. F. The E. M. F. of the latter therefore opposes the current.

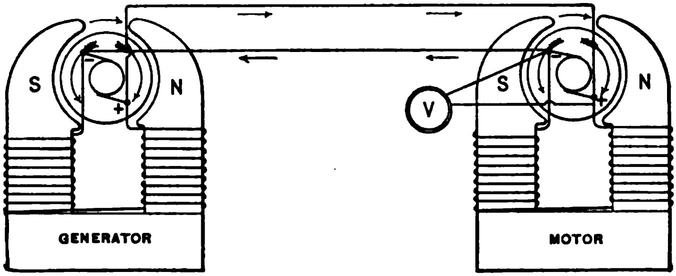


Fig. 408

If the motor is provided with a fly wheel to keep up its motion when the generator current is cut off, a voltmeter placed across its terminals, as V in the figure, will show only a slightly diminished E. M. F. immediately after the circuit is broken, if there is no load on the motor to produce a quick slackening of speed. The voltmeter indicates no reversal of current when the generator is cut off. This fact shows that the positive brush of the generator is connected to the positive of the motor, or that the E. M. F. of the motor is a back E. M. F. The voltmeter may be replaced by an incandescent lamp; it will glow nearly as brightly for a few seconds directly after the main circuit is opened as before.

642. Relation of Speed to Field Strength. — It has already been demonstrated that the work done against an opposing E. M. F. per second is measured electrically by the product of this counter E. M. F. and the current, or $E'I$ (§ 565).

The two factors of the power, measured mechanically, are the *torque* and the *speed* (that is, torque and angular velocity, or $T\omega$). The torque is the moment of the couple producing the rotation; it is proportional both to the strength of the field and the current in the armature. If the field is kept constant, the torque is proportional to the current; the back E. M. F. is proportional to the speed. Hence we may write,

$$E'I = AnT, \tag{115}$$

where T is the torque, n the number of rotations per second, and A a constant or proportionality factor ($A = 2\pi$).

When the motor is working under a fixed load, an increase in the field strength increases the torque and decreases the speed n; weakening the field, on the other hand, diminishes the torque and increases the speed. Both these conclusions follow from the constancy of the product nT under the assumed condition of the fixed load. In both cases the speed changes until the counter E. M. F. acquires the same value that it had before the change was made in the field strength.

II. ALTERNATING CURRENT MACHINES

643. Alternators. — If the brushes of a dynamo bear on two continuous rings on the shaft, instead of a commutator, the current in the external circuit will alternate or reverse every time the armature turns through the angular distance from one field pole to the next. A complete series of changes in the current and the E. M. F. takes place while the armature is turning the angular distance of one pole to the next one of the same sign. This series of changes is called a *cycle*. The frequency, or number of cycles per second, is the product of the number of *pairs of poles* of the field magnet and the number of rotations per second. Frequencies are now restricted to the limits of about twenty-five and sixty cycles per second. Multipolar machines are used to avoid excessive speed of rotation.

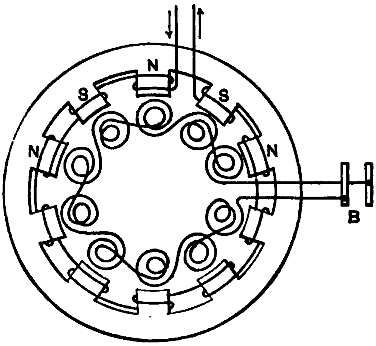


Fig. 409

Figure 409 is a diagrammatic sketch of an alternator with a stationary field and a rotating armature. In large ma-

chines the armature is usually the stationary member and the field the revolving one. The field is excited by a direct current machine. The armature coils are reversed in winding from pole to pole; they are all joined in series and the terminals are brought out to the two slip rings at *B*. The brushes bearing on these rings lead to the external circuit.

644. Lag of the Current behind the Electromotive Force. — — When an alternating electromotive force is applied to a circuit having inductance, one of the novel and essential facts is that the current reaches its maximum value later than the electromotive force; and, as a consequence, Ohm's law is no longer adequate to express its value. The effect of self-inductance is not only to introduce an additional electromotive force, but to produce a lag of the current in phase behind the electromotive force impressed on the circuit by the generator.

The effort required to get an understanding of the phase lag of the current is well worth while. Let an alternating current, following the simple harmonic law, be represented by the heavy sine curve I in Figure 410. The induced electro-

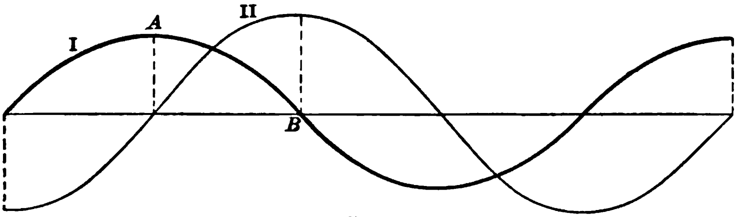


Fig. 410

motive force is proportional to the rate of change of magnetic flux, and hence to the rate of change of the current producing it when there is no iron in or about the circuit. The induced electromotive force curve is represented by the light line II. This is also a sine curve, for the rate of variation of a sine function is itself a sine function. The curve of induced E. M. F. reaches its maximum value a quarter

of a period *later* than the current curve; for, when the current is a maximum at A, its rate of change is zero and the induced E. M. F. is zero; when it decreases through its zero value at B, its rate of change is a maximum, and the induced E. M. F. is a maximum in the *positive* direction. Hence the induced E. M. F. is a quarter of a period behind the current, or the current and induced electromotive force are said to be in quadrature.

The electromotive force producing the current by Ohm's law corresponds in phase with the current itself. The electromotive forces consumed by resistance and self-inductance may therefore be represented by the two adjacent sides of a right triangle (Fig. 411), where ab and bc are the electromotive forces consumed by resistance and inductance respectively. If these electromotive forces are added geometrically as if they were vector quantities, their resultant is the hypotenuse ac, which is therefore the maximum impressed E. M. F. But the current agrees in phase with ab; it therefore lags behind the impressed electromotive force by the angle ϕ. In the absence of capacity in the circuit, this angle of lag becomes zero only when the inductance is zero.

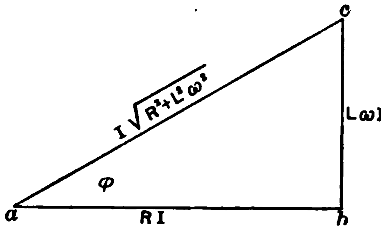


Fig. 411

One component of the impressed E. M. F. goes to neutralize the E. M. F. of self-induction; the other is lost in producing the current against resistance by Ohm's law.

Figure 412 A was reproduced from one of Professor Thomas's oscillograms and shows the current lagging behind the E. M. F. nearly a quarter of a period in a circuit of large self-inductance. Self-inductance operates to check sudden changes in the impressed E. M. F. Hence the small inequalities of the E. M. F. curve, due to the presence of higher harmonics or multiples of the fundamental frequency, nearly or quite disappear in the current curve.

The effects of *capacity* in a circuit are exactly the opposite of those of self-induction. Capacity causes the current to lead the impressed E. M. F.

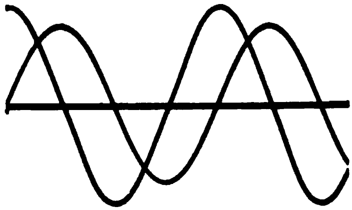


Fig. 412 A

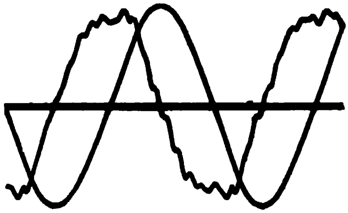


Fig. 412 B

and it exaggerates the higher harmonics instead of reducing or suppressing them in the current (Fig. 412 *B*). Figure 412 *C* is an oscillogram from the same machine, but in this case the circuit was non-inductive and without capacity. The current and E. M. F. are in the same phase and the current at the maximum reproduces all the small irregularities of the E. M. F. due to the presence of odd harmonics.

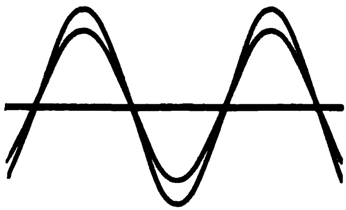


Fig. 412 C

645. Value of an Alternating Current. — The maximum value of the E. M. F. of self-induction is $L\omega I$, where L is the coefficient of self-induction and ω the angular velocity, or $2\pi n$. The impressed E. M. F. represented by the hypotenuse of the triangle is $I(R^2 + L^2\omega^2)^{\frac{1}{2}}$. Hence

$$E = I(R^2 + L^2\omega^2)^{\frac{1}{2}}, \text{ and } I = E/(R^2 + L^2\omega^2)^{\frac{1}{2}}. \quad (116)$$

The expression $(R^2 + L^2\omega^2)^{\frac{1}{2}}$ is called the *impedance* of the circuit. It shows that the effect of inductance on the value of the current is equivalent to an additional resistance.

646. Effective Volts and Amperes. — The heating effect of an alternating current, as in the filament of an incandescent lamp, is proportional to its mean square value, and all instruments for measuring alternating currents and pressures indicate "square root of mean square" values and not the

arithmetical mean. Thus the electrodynamometer (§ 593), whether used to measure current or voltage, integrates the forces operating it, and these are proportional either to the squares of the current or of the electric pressure. If the current and the electromotive force follow the sine law, the mean given by the measuring instruments is 0.707 of the maximum value. When a voltmeter on an alternating circuit reads 70.7, the voltage alternately rises to + 100 and sinks to − 100 as positive and negative maxima. The values given by measuring instruments for alternating current circuits are *effective volts* and *effective amperes*.

647. **Transformers.** — A *transformer* is an induction coil with a primary of many turns, a secondary of a smaller number, and a closed magnetic circuit. It is employed with alternating currents as a "step-down" instrument for the purpose of reducing the high electromotive force on the transmission line to a low electromotive force for lighting and power. It is entirely reversible and can be used equally well for the "step-up" process with alternating currents.

The primary and the secondary coil are both wound around the same iron core, but are insulated from each other as perfectly as possible. The iron (Fig. 413) serves as a path for the flux of magnetic induction. The relation of the flux and the current is a reciprocal one, so that they may always exchange places. In the "shell type" of transformer the iron surrounds the wire, while in the "core type" the wire surrounds the iron. With either relative arrangement of the iron and the coils, nearly all the lines of induction produced by the primary pass also through the secondary, and *vice versa*.

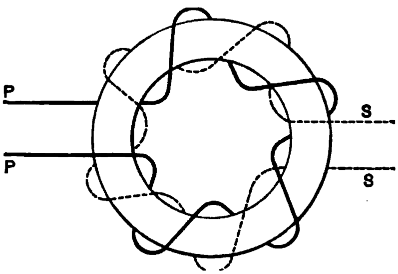


Fig. 413

When the secondary is open the primary acts simply as a "choke coil," and the only current passing is the very small one required to magnetize the iron for the generation of a counter E. M. F. of self-induction nearly equal to the impressed E. M. F., and to furnish the small amount of energy lost in hysteresis. The electromotive forces generated in the two coils are directly as the ratio of the number of turns on the two; while the currents in them when the secondary is closed are very nearly in the inverse ratio of the number of turns. The energy in the secondary circuit is nearly the same as that expended on the primary. The small loss is chargeable to loss in the copper of the primary and to losses in heating the iron on account of hysteresis and eddy currents.

The secondary current is nearly opposite in phase to the primary, and causes a diminution in the apparent self-inductance of the primary coil; so that the larger the secondary current, the larger the primary. A transformer is thus nearly self-governing.

648. Polyphase Currents. — It has long been known that two or more alternating currents of the same frequency, but differing in phase, may be obtained from one generator. If, instead of a commutator, four insulated rings on the shaft be connected to four equidistant points on either a drum armature or a Gramme ring, the currents in the externally separate circuits will differ in phase by a quarter of a period. Each pair of rings connected to points on the armature 180° apart act as terminals for an alternating circuit. It is obvious that one current passes through its maximum at the instant that the other

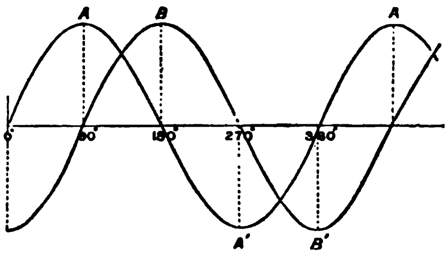


Fig. 414

passes through its zero value (Fig. 414). In a similar way, three-phase currents may be taken from rings as terminals if they are connected to conductors in the armature 120° apart. If there are only three conductors, each one serves as a return for the other two, since the algebraic sum of the two currents having the same sign at any instant is always equal and opposite to the third (Fig. 415).

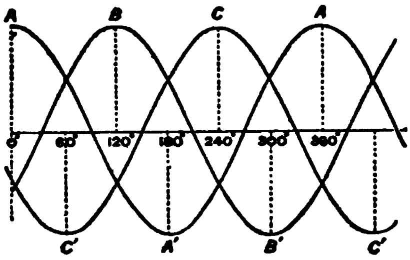


Fig. 415

649. The Rotatory Field. — When an alternating current is transmitted through a coil without iron, it produces an alternating magnetic field along the axis of the coil. If the current follows the sine law, the magnetic flux will follow the sine law also.

Let two such coils be set with their axes at right angles, and let equal two-phase alternating currents pass through them. Two simple harmonic motions of equal amplitude, at right angles, and differing in phase by a quarter of a period, combine to produce uniform circular motion (§ 38). In a similar way the two coils *AA* and *BB* (Fig. 416) produce a rotating magnetic field near their common center. In 1888 Ferraris in Italy mounted within them a hollow copper cylinder on pivots at top and bottom. When two-phase currents are passed through the two circuits of the Ferraris apparatus, the copper cylinder is set rotating in the direction of the rotating field. The rotation of the field produces currents in the copper, as in Arago's rotations. By Lenz's law the motion

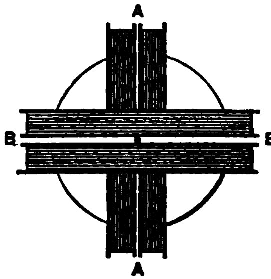


Fig. 416

of the cylinder is in the direction to check the induction in it; hence the cylinder is dragged around in the same direction as the rotation of the field; for, if the speed of the cylinder were the same as that of the field, no current would be induced in the copper.

If one current is reversed with respect to the other, that is, if its phase is changed by 180°, the direction of rotation of both the field and the cylinder is reversed. The cylinder tends to run up to synchronism with the field, but never quite reaches it; the difference in their speeds, called the *slip*, is just sufficient to induce currents to supply the requisite torque. If the rotation of the field generates a direct electromotive force, the rotation of the cylinder, which is equivalent to the rotation of the field in the other direction, generates a counter electromotive force, and the latter is always smaller than the former.

650. The Induction Motor. — Rotation of the magnetic field may also be secured by winding the coils for the two circuits on an iron ring (Fig. 417). The coils A and A' are wound so as to make consequent poles at B and B', while the coils B and B' are wound to produce consequent poles at A and A'.

At the instant when the current through coils A and A' in phase A has reached its maximum value at A (Fig. 414), the current in phase B is zero and the poles in the ring are concentrated at B and B'. An eighth of a period, or 45°, later the currents in the two phases are positive and equal. The poles in the ring then extend from A to B and from A' to B'. An eighth of a period later, or at B in phase B (Fig. 414), the current in phase A is zero while the current in phase B is a maximum. The poles are now at A and A', and they have moved forward counter clockwise a quarter of the way around the ring.

The rotation of the field within the ring of Figure 417 may readily be shown by placing a glass vessel in the horizontal ring, filling with water

to the level of the middle of the ring, and sifting powdered iron on the surface of the water. With small two-phase currents traversing the coils, the individual iron particles rotate rapidly; with larger currents, the whole surface rotates in the direction of the rotation of the field.

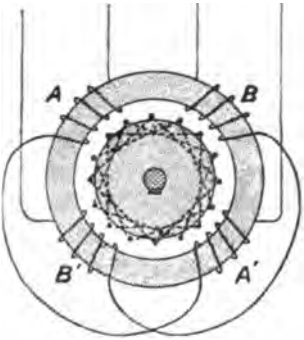


Fig. 417

The Δ connections for a three-phase winding of the stationary member, or *stator*, of an induction motor are shown in Figure 418. Instead of three coils, P, Q, and R, there may be any multiple of three.

Inside the two-phase or the three-phase ring is mounted a *rotor*, consisting of a laminated iron cylinder with heavy conductors embedded in the periphery parallel to its axis of rotation. They are connected together at the ends of the cylinder so as to form a "squirrel cage" of copper. The induced currents in this cage produce a torque which turns the cylinder in the direction of the rotating field.

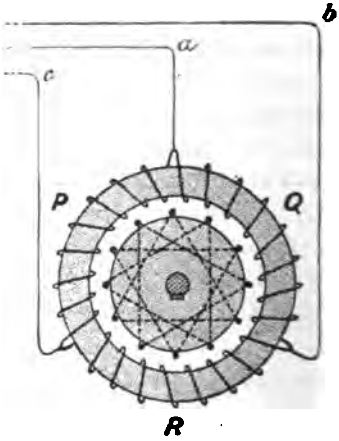


Fig. 418

651. Power in an Alternating Current Circuit. — It has been shown that for a direct constant current the energy per second is the product EI of the E. M. F. and current (§ 536). Similarly the power of an alternating current is the product ei of corresponding instantaneous values of the E. M. F. and current. But this product is incessantly changing; it is instructive to resort to the graphical method to illustrate the changes taking place during a few cycles.

In Figure 419 the curve of current I in an inductive circuit is behind the electromotive force curve E. If we take the product ei for successive values for several periods, and

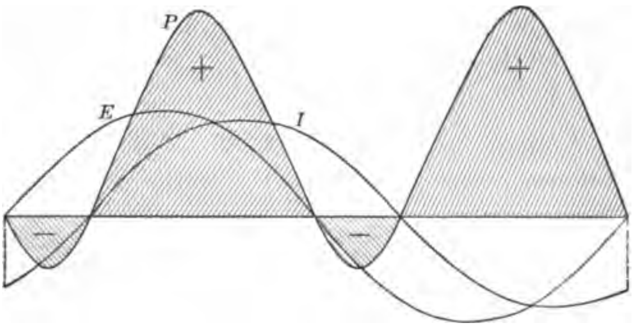


Fig. 419

plot them with their proper signs, the result will be the power curve P, which is also harmonic if the other curves are.

It will be noticed that the frequency of the power curve is double that of the E. M. F. and current curves, and that half the areas are positive and half are negative. The interpretation of the signs is that during the intervals of time covered by the positive products the generator is putting energy into the circuit; while during the intervals covered by the negative areas, the circuit is returning energy to the generator. In other words, there is a give-and-take of energy by the inductive circuit similar to the mechanical action of a flywheel.

The power transmitted is represented by the difference between the positive and negative areas. When there is no phase difference between the current and the electromotive force, all the areas are positive, and the power transmitted is a maximum for the given E. M. F. and current; with a phase difference of a quarter of a period, the negative areas are equal to the positive and no power is transmitted. For other phase differences the power is intermediate, and varies

as the cosine of the angle of lag. If E and I are the effective E. M. F. and current, the power transmitted is $EI \cos \phi$, in which ϕ is the angle of lag. The term $\cos \phi$ is called the *power factor*. Even when the E. M. F. and current are not simple harmonic, the factor by which the product EI must be multiplied to give the power transmitted is still called the power factor.

Problems

1. If the temperature coefficient for the resistivity of copper is 0.0039, and that for a carbon filament − 0.0003, how many ohms of copper resistance must be joined in series with a carbon filament of 229 ohms, so that the combined resistance may be constant while both undergo equal changes of temperature?

2. If a dynamo machine generates 85 kilowatts of electric power and the gross efficiency of the machine is 85 per cent, what horse power is required to drive it?

3. A building is lighted with 1000 incandescent lamps, each taking 0.25 ampere at 110 volts. Allowing a loss of energy in the mains of 5 per cent, and in the dynamo 15 per cent, how many kilowatts are applied to run the machine?

4. A coil of wire is moved across a magnetic field in such a way that it cuts lines of magnetic force at the rate of 5×10^9 per second. What is the average electromotive force generated?

5. A rectangular loop of wire 20 cm. wide and 30 cm. long is mounted symmetrically on an axis parallel to its length and rotates uniformly at the rate of 1200 revolutions per minute across a uniform magnetic field of 5000 lines per square cm. Find the average electromotive force in volts generated.

6. The armature of a two-pole dynamo machine is wound with 480 conductors on its surface; the total magnetic flux through it is 1,250,000 lines; compute the E. M. F. in volts when it rotates at the rate of 1200 revolutions per minute.

7. A two-pole electric motor is wound with 128 wires on the outside of the armature; the total magnetic flux through it is 1,250,000 lines. Find the work done in ergs in one revolution when a current of 50 amperes flows through each wire; also find the power in kilowatts when there are 960 revolutions per minute.

CHAPTER XXII

ELECTRIC OSCILLATIONS AND WAVES

I. ELECTRIC OSCILLATIONS

652. Oscillatory Discharges. — The oscillatory discharge of a Leyden jar has already been described (§ 521). The discharge of any condenser through a discharge circuit of low resistance is oscillatory. The charge has the appearance of possessing inertia; when a condenser is suddenly discharged through a low resistance, the first rush surges beyond the condition of equilibrium, and the condenser is charged in the opposite sense; a reverse discharge follows, and so on, — each successive pulse being weaker than the preceding, until after a few surges the oscillations cease. These alternating surges of high frequency are called *electric oscillations*.

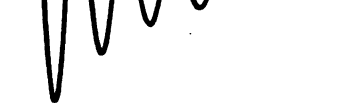


Fig. 420

A beautiful instance of the oscillatory discharge of a condenser is the photographic reproduction of one of Professor Thomas's oscillograms (Fig. 420). The damping out of such a discharge is shown by the rapid decrease in the amplitude of the successive discharges.

The oscillations of a small Leyden jar, charged by connecting its coatings with the secondary terminals of an induction coil, may readily be exhibited to a large number of persons. It is convenient, though not

essential, to close and open the primary circuit by means of a seconds pendulum. A pointed strip of tin foil is brought over from the inner coating of the jar, so as to leave a small spark gap between it and a similar pointed strip connected with the outer coating. At every break of the primary circuit a spark will leap across this small gap if the adjustments are properly made. If it is viewed in a four-square mirror, rotating with moderate speed, each principal discharge will be found to consist of from about four to twelve successive images. A single observer may view it in a telescope after reflection from a plane mirror on the end of a tuning fork making about one hundred vibrations a second. The rate of oscillation in this case is comparatively low on account of the large self-inductance of the secondary coil, but the whole series of oscillations takes place in the "incredibly short space of time occupied by a spark."

653. Period of an Oscillation. — Whether the discharge of a condenser is unidirectional or oscillatory depends on the relation between the resistance and self-inductance of the discharge circuit and the capacity of the condenser. When a condenser is discharged through a low resistance, oscillations take place because the choking reactions due to self-induction are neutralized by the capacity. The oscillations then continue, like the vibrations of a tuning fork, until their energy is expended partly in heat and partly in a manner to be described directly.

The condition required for capacity to neutralize self-induction is expressed in the form of the equation $L\omega = 1/C\omega$, where C is the capacity of the condenser. Since $\omega = 2\pi n$ and the period $T = 1/n$, if we solve this equation of condition for T, we get

$$T = 2\pi\sqrt{CL}. \tag{117}$$

The period of an oscillation may therefore be adjusted by changing either the capacity or the self-inductance of the circuit.

Martienssen has described an oscillating system of colossal proportions. A huge capacity, consisting of 500 two-microfarad condensers in parallel, was combined with a self-inductance of 1000 henrys. The resistance of the discharge circuit was 50 ohms. Applying formula (117), the period of oscillation is 6.28 seconds. A direct current am-

meter in circuit followed this slow oscillation and indicated five complete oscillations in the discharge.

654. Visible Demonstration of Electrical Resonance.—Electrical resonance in coils of wire wound on glass or wood may be made visible by the following experiments due to Seibt:

The connections and means of adjustment for varying the frequency are shown in Figure 421. *J* is an induction coil capable of giving a spark at least 20 cm. long; C_1 and C_2 are Leyden jars, which may be joined either in series or in parallel for the purpose of varying the capacity; *F* is an adjustable spark gap between zinc balls, and the discharge circuit of the jars is through the adjustable inductances L_1 and L_2; *R* is a resonance coil connected at the bottom to the point *P* of the discharge circuit. The adjustable inductances consist of stout copper wires wound on porcelain insulators.

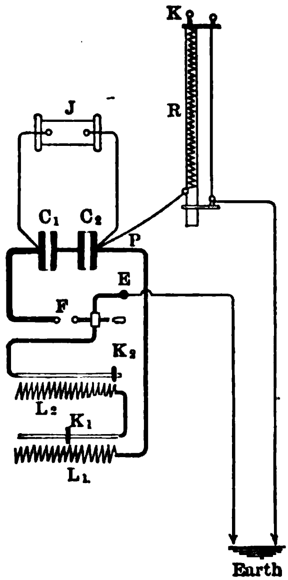


Fig. 421

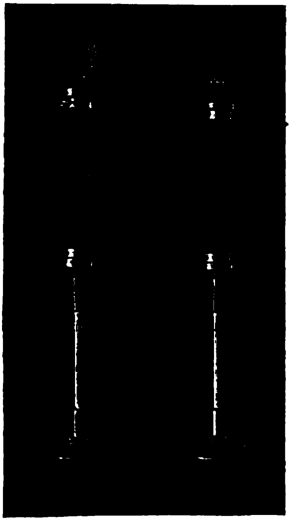


Fig. 422

The fundamental phenomenon of resonance is demonstrated by means of two coils about 30 cm. long, but wound with a different number of turns of wire (Fig. 422). When the two are connected to the point *P*, and the self-inductance is adjusted by means of the sliding contacts K_1 and K_2, resonance is indicated by one of the coils showing a brilliant violet brush discharge from the knob at the upper end, while the other coil remains dark. By changing the period of oscillation, the second coil becomes luminous and the first one remains dark. The wire terminates at the knob on top; and when the electrical

oscillations in the coil are in unison with those of the discharge circuit of the Leyden jars, the tube is the analogue of a closed organ pipe, giving its fundamental tone. There is a node at the upper end where the change in the electric pressure, due to the surges of the charge, is the greatest. The frequencies of oscillation for the two tubes are not the same.

Fig. 423

Figure 423 represents the appearance in the dark of a coil two meters long. Parallel with the coil is stretched a fine steel wire; its lower end is connected with a gas or water pipe, and the upper end is insulated from the coil. When the discharge circuit is properly adjusted, a brush discharge extends from the node at the top part way down the coil, gradually fading out. By increasing the frequency of the oscillations, the coil may be divided into two segments with an additional node one third of the distance from the lower end. With the steel wire closer to the coil, the discharge circuit may be adjusted so as to show the coil divided into three, four, or even more segments. The vibration frequencies are then 3, 5, and 7 times the fundamental.

Fig. 424

If the steel wire be joined to the coil at the upper end, the arrangement then represents the conditions obtaining in an open organ pipe, the charges surging freely in and out at both ends. For the fundamental frequency the node is now at the middle, as shown in Figure 424. For the first overtone there are two nodes, at one quarter of the length of the coil from each end. Higher overtones may also be obtained in this case.

The nodes and antinodes may also be made visible by carrying a small Geissler tube along the coil. It will light up brightly at the nodes, but will remain dark at the antinodes.

A coil of half the length of Figure 424, but with the same number of turns per centimeter, and with the upper end not connected with the steel wire, vibrates in unison with the coil of double the length and

connected to the ground at the upper end. These two coils are the analogues of an open organ pipe and a closed one of half the length. The frequency is the same.

655. The Singing Arc. — When a condenser and an impedance coil are connected in series around an electric arc supplied with direct current, the arc itself may be thrown into a state of vibration with a frequency depending on the period of the shunt circuit (Fig. 425). Small currents of 1 to 3 amperes give the best results.

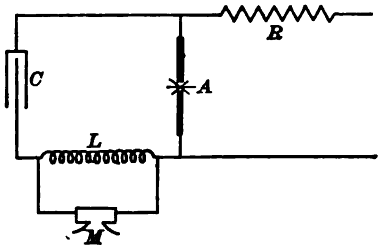


Fig. 425

Sudden changes occur in the resistance of the arc, and these change the potential difference between the carbons available for charging the condenser. The oscillations thus started continue with the natural period of the oscillating system, including the arc itself. The frequencies are lower than those of the condenser circuit without the arc, indicating that the arc acts as if it had self-inductance. With graphite electrodes the frequency may be as high as several hundred thousand a second.

In air compressed to six atmospheres the oscillating arc, with a current of 0.2 ampere and a pressure of 4500 volts, has all the characteristics of a rapid spark discharge.

With a suitable microphone transmitter M connected as a shunt to the impedance L in the condenser circuit, the arc may be made to serve as a loud-speaking telephone receiver. For this purpose a long arc, lengthened by means of an impregnated positive carbon, gives the best results. Numbers, words, and phrases are repeated with weird effect.

II. ELECTROMAGNETIC WAVES

656. Henry's Experiment. — When an electric current rises from zero to its maximum value in a conductor, a magnetic field is set up about it; when the current is increased, the magnetic lines enlarge and new ones push out from the conductor. When the circuit is opened or the current is reversed, these lines close in on the conductor and restore to it the energy stored in the ether through an E.M.F. of self-induction. If the current oscillates with great rapidity, part of the energy radiates into space in the form of electromagnetic waves in the surrounding medium. With the slow alternations employed commercially, the loss by radiation is inappreciable; but such is no longer the case when the rate of oscillation is a million or more per second.

Joseph Henry appears to have been the first to detect electromagnetic waves radiating from a circuit running around a room, when an inch spark from the prime conductor of a frictional machine was thrown on to the end of the circuit. Sewing needles were magnetized in a parallel circuit thirty feet below, with two floors and a ceiling intervening. He says, "The diffusion of motion in this case is almost comparable with that of a spark from a flint and steel in the case of light." Thanks to the remarkable researches of Hertz, we now know it to be the same. The magnetic field produced by the discharge through the one conductor spread with the velocity of light to the closed circuit below, where a part of its energy was absorbed by cutting through the circuit, and resulted in an electric flow sufficient to magnetize the needles placed in a helix forming a part of the circuit.

The energy stored in a Leyden jar is not all dissipated in the heat of the spark, but some of it is radiated into space in the form of electrostatic and electromagnetic waves.

657. Hertz's Researches. — To Hertz belongs the credit of having put the theory of electromagnetic waves to the test of experiment. The simplicity of his appliances is no

less remarkable than the novelty and importance of the results derived from them. With the insight of genius he seized on the only available means of producing electric waves short enough to be readily measured, viz., the disturbances propagated outward from the discharge of a condenser of small capacity.

Hertz's apparatus to serve as the source of waves he called an *oscillator* (Fig. 426). It consisted of two metallic

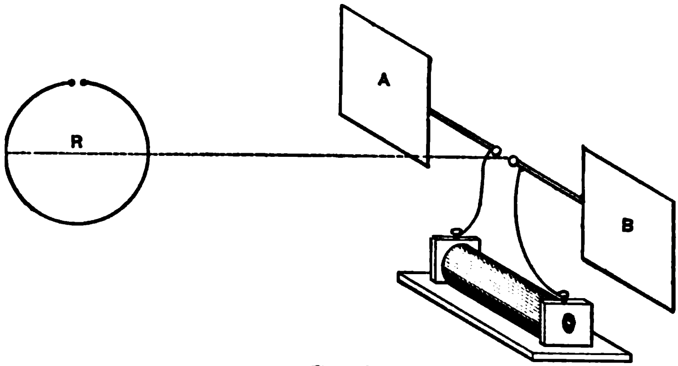


Fig. 426

plates A and B, 40 cm. square and mounted 60 cm. apart. The balls at the spark gap were polished. The receiver, or *resonator*, was a circle of wire 70 cm. in diameter, and its spark gap was adjustable by means of a micrometer screw. The two plates of the oscillator were connected to the terminals of the secondary of an induction coil. The plates formed a condenser of small capacity, and the discharge across from ball to ball was oscillatory. The frequency was definite, and a succession of electrostatic and electromagnetic waves were emitted with a half period of $\frac{1}{100000000}$ sec.

The finite speed of the waves was demonstrated by placing a large plane sheet of zinc on a distant wall of the room and observing the sparks produced at the minute break in the tuned resonator at different positions along the dotted line. The sheet metal acted as a reflector, so that stationary waves

were produced by interference between the direct and reflected systems precisely as in sound (§ 197). The nodes and antinodes were located with considerable precision. The distance between them determined the wave length, and the frequency of the oscillations gave the velocity of transmission. This was found to be of the same order of magnitude as the known velocity of light.

By the aid of large parabolic zinc reflectors Hertz demonstrated that electric waves are reflected to a focus in the same manner as light. He also constructed a huge prism of asphaltum, which is transparent to electric waves, and measured its index of refraction. Gratings consisting of parallel conducting wires exhibited polarization phenomena.

Thus Hertz demonstrated that the waves radiating from an oscillatory spark discharge are capable of reflection, refraction, and polarization the same as light. They possess all the characteristics of light except in point of wave length.

658. The Coherer. — One of the very sensitive instruments for detecting electromagnetic waves is called a *coherer*. In 1900 Branly discovered that loosely packed metal filings, which offer almost infinite resistance to an electric current, acquire good conductivity instantly under the influence of electric waves. They lose this conductivity when they are shaken or lightly tapped.

Fig. 427

A mixture of 95 per cent nickel and 5 per cent silver in an exhausted glass tube has been successfully used by Marconi (Fig. 427). The electrodes S and S_1 are silver and between them is a small quantity of the loosely packed filings.

An imperceptible discharge through a coherer from the cover of an electrophorus is quite as effective as electric waves in lowering the resistance of the filings. The delicacy of the coherer as a detector is so great that it responds even when the source of powerful electric oscillations is scores of miles away. On account of this sensitiveness, it must be screened from all minute electric waves produced by the apparatus in

the local voltaic circuit used for receiving, by inclosing it in a metallic box.

The passage of electric waves through a coherer probably causes minute sparks across from filing to filing; these either break down the contact resistance due to condensed gases, or they may even make slight welds between adjacent particles. A current from one or two voltaic cells then readily follows.

659. Wireless Telegraphy. — The transmission of electric oscillations for laboratory purposes may be accomplished by means of apparatus

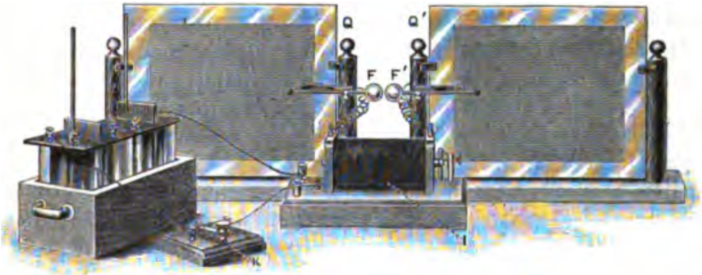


Fig. 428

similar to that represented in Figure 428. Two sheets of tin foil on glass Q and Q' are connected with the two discharge balls F and F' of an induction coil.

The *receiver* tuned to the same frequency as the transmitter is somewhat more complicated, but the different parts and the connections may

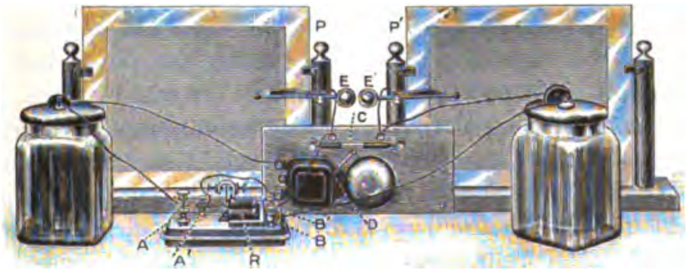


Fig. 429

be made out in Figure 429. The metallic plates P and P' are duplicates of those in the transmitter. The balls E and E' are connected both with the plates and with the ends of coherer C. A circuit is formed,

including the voltaic cell on the right, the coherer, and the fine wire coil *R* of a relay through the binding posts *B* and *B'*. Another circuit is formed through the contact points of the relay, the voltaic cell at the left, and the vibrating bell. The bell is so placed that its hammer strikes the coherer when the bell rings. The transmitter and receiver may be placed at some distance from each other. When the key *K* of the transmitter is firmly pressed down for a moment, a spark passes across the gap between the balls; at the same instant the coherer allows a current to pass through, and this current operates the relay and closes the circuit through the electric bell. The blow of the hammer on the coherer restores it to its sensitive condition of high resistance, ready for the reception of another short succession of electric oscillations.

Marconi has succeeded in transmitting wireless messages over many hundreds of miles by using very powerful oscillators and extremely sensitive detectors. This system has been particularly successful in communications between ships at sea, and has been instrumental in averting some frightful calamities. Ships fitted with Marconi's system maintain communication when one hundred and fifty or two hundred miles apart at sea. Instead of plates, tall masts and parallel wires are used in connection with both the transmitter and the receiver. Auxiliary devices are employed which cannot be described here.

III. SOME RELATIONS BETWEEN LIGHT AND MAGNETISM

660. Rotation by a Magnetic Field. — The first definite relation between light and magnetism was discovered by Faraday in 1845. When a beam of plane polarized light is transmitted through a refractive medium and parallel to a field of magnetic force, the plane of vibration is rotated. The effect is especially marked in carbon bisulphide and in dense flint glass, but it is very feeble in the case of gases.

A beam of light, polarized by transmission through a Nicol's prism, is passed through a prism of heavy glass (borosilicate of lead), with parallel polished ends, and placed in a powerful magnetic field, the direction of which coincides with that of the beam of light. A second Nicol's prism as an analyzer receives the beam; it is first turned so as to cut off all the light from the polarizer.

When the magnet is excited light passes through the analyzer. It may be extinguished by rotating the analyzer

through a small angle, but it is not possible to get complete extinction for white light; colors appear, showing that the angle of rotation is a function of the wave length. The rotation is nearly inversely as the square of the wave length. If the electromagnet is large, it will be apparent that time is required to magnetize it, inasmuch as the transmitted light grows sensibly in intensity for a second or more after closing the circuit through the coils of the magnet. On the other hand, Sir Oliver Lodge has shown that the rotation of the beam of light, first in one direction and then in the other, follows the oscillations of the discharge of a Leyden jar through coils producing a magnetic field without iron.

For a given wave length the rotation is proportional to the field intensity and to the thickness of the refractive medium. The rotation produced by a thickness of one centimeter in a field of unit intensity for *D* light is: for water, 0.0131°; for carbon bisulphide, 0.0435°; for benzene, 0.0297°; for dense flint glass, 0.06°. Very large rotations are produced by thin semitransparent films of iron and other magnetic substances in a magnetic field.

When the plane of vibration in a transparent substance is rotated by magnetism, the direction of rotation is reversed with the reversal of the field; if therefore the beam is reflected back through the medium in the same field, the rotation is doubled. Advantage has been taken of this fact to increase the angle of rotation.

661. The Zeeman Effect. — Faraday examined the spectra of light from a source placed in a magnetic field, with the object of finding an effect of magnetism on radiation. The result was negative. In 1896 Zeeman repeated Faraday's experiments with a Rowland grating of high resolving power and discovered an effect which has proved to be important in relation to theories of radiation, as well as in its application to magnetic phenomena.

The source of light must be in a strong magnetic field,

such as that between the poles of a powerful electromagnet. When light from such a source travels either at right angles to the direction of the field or parallel to it, the spectral lines become double, triple, or in some cases still more complex. Lorentz predicted that this modified radiation would be polarized, either plane or circularly, according to the direction in which it is viewed relative to the magnetic field. This prediction proved to be true.

In the simplest case of a source viewed at right angles to the lines of the field, each spectral line becomes a triplet, the vibrations of the middle undisplaced component being parallel to the lines of force, and those of the lateral components at right angles to them.

When the source is viewed in a direction along the lines of force, single spectral lines are doubled, the two components being displaced in opposite directions, and both of them circularly polarized.

Lorentz and Lorimer have shown that the Zeeman effect can be accounted for if we assume the radiant center in the case of a line spectrum to be a minute electrically charged particle, called an *electron* (§ 666), rotating about an atom in a circular or elliptical orbit, and that these electrons communicate their motion to the ether. A charge of electricity on a moving electron is equivalent to an electric current, and has a magnetic field which may react with any external field. When these electrons move in a magnetic field, their natural periods are subject to perturbations, and these perturbations may account for the changes in the period of vibration indicated by the displacement of the resolved lines.

CHAPTER XXIII

PASSAGE OF ELECTRICITY THROUGH GASES

I. DISCHARGES THROUGH VACUA

662. Gases as Conductors. — All gases under normal conditions of temperature and pressure are such extremely poor conductors of electricity that it is very difficult to measure their conductance. When a potential difference is established between two points in a gas, the gas is put under stress, as already described in §§ 526, 527. The stress increases with the potential difference until the gas breaks down as an insulator and a discharge passes. To produce such a discharge, or momentary current, a high potential is necessary, — several thousand volts for a spark 1 cm. long in air at atmospheric pressure and temperature. The necessary electric pressure depends on the shape of the electrodes and on the nature, pressure, and temperature of the gas. The surface density and the stress in the gas at a sharp point are greater than at a plane surface; hence the power of points to produce the glow discharge, sometimes known as Saint Elmo's fire.

663. Effects of reducing the Pressure. — When the air is gradually exhausted from a glass tube, provided with metal electrodes at its ends, an electric discharge passes with greater and greater ease until a definite minimum pressure is reached. If the exhaustion is carried still further, the voltage required to produce a discharge increases somewhat rapidly, and at the lowest attainable pressure no discharge can be made to pass. The critical pressure is that at which a

minimum potential difference is required to cause a discharge. It varies with the distance between the electrodes.

The appearance of the discharge changes in a marked manner as the exhaustion proceeds. When a pressure of about half a millimeter of mercury is reached, the general appearance of the tube when a discharge passes is imperfectly shown in Figure 430. At the cathode *C* is a soft glow which moves

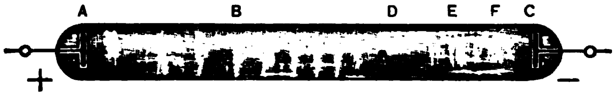


Fig. 430

about over the flat electrode. Next to this is the first dark space *F*, called the Crookes dark space. It is comparatively non-luminous. Immediately beyond the Crookes dark space is a luminous region *E* called the negative column or glow. Still further along is a second negative dark region *D*. Between this and the anode *A* is a luminous portion *B* extending up to the anode and known as the positive column.

When the length of the tube is increased beyond a few centimeters, the negative column and dark space are confined to the neighborhood of the cathode, but the positive column stretches away to the anode.

The luminosity of the positive column is not usually continuous, but is composed of alternate dark and light spaces called *striæ*. Sir J. J. Thomson excited an exhausted tube 50 feet long; with the exception of a few inches near the cathode, the positive column filled the whole of the tube, with marked striations throughout its length. The color of the striæ depends on the nature of the residual gas. They present a peculiar unstable flickering motion, similar to that sometimes observed during an auroral display.

The celebrity of the tubes made many years ago by Geissler gave to them the name Geissler tubes. Some of the patterns are shown in Figure 431. The narrow portions of a tube, containing hydrogen or vapor of water, glow with a brilliant crimson. Fluorescent materials in Geiss-

ler tubes are beautifully luminous. Uranium glass, and solutions of quinine, æsculin, and naphthalene-red in tubes surrounding the exhausted one are among the best examples of fluorescence.

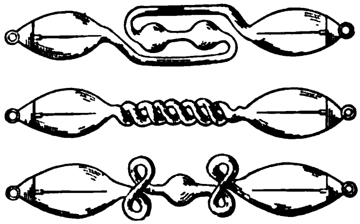


Fig. 431

664. Cathode Rays. — When the pressure in a tube is reduced below a thousandth of a millimeter of mercury, the character of the discharge is much altered; the positive column gradually disappears, and the sides of the tube glow with brilliant phosphorescence. With English glass the glow is blue; with German glass it is a soft emerald. The phosphorescence is produced by something radiated in straight lines from the cathode and known as *cathode rays*. They have been studied at length by Sir William Crookes, and tubes for the purpose are called *Crookes tubes*.

Fig. 432

Many other substances besides glass are caused to glow by the impact of cathode rays (Fig. 432), such as diamond, ruby, and various sulphides. The color of the glow depends on the substance.

Fig. 433

Cathode rays, unlike rays of light, are deflected by a magnet, and when once deflected they do not regain their former direction (Fig. 433),

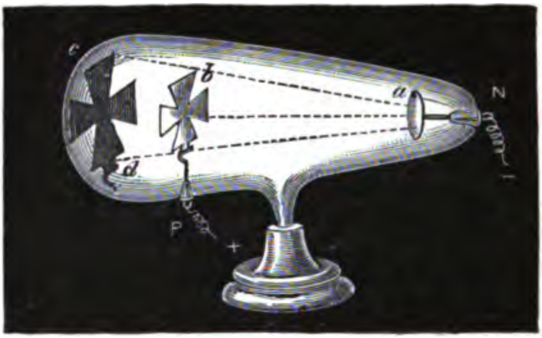


Fig. 434

differing in this respect from the discharge stream or current in a Geissler tube. Cathode rays proceed in straight lines, except as they are deflected by a magnet or by mutual repulsion. A screen placed across their path intercepts them and casts an apparent shadow on the walls of the tube; there is no phosphorescence within this geometrical shadow (Fig. 434). If the cathode, or any obstruction on which the cathode rays fall, is delicately poised, it will be set in motion by this electrical stream or wind.

Fig. 435

When the cathode is made in the form of a concave cup, the rays are brought to a focus at its center of curvature; platinum foil placed at this focus is raised to bright incandescence, and may be fused (Fig. 435). Glass on which an energetic cathode stream falls may be heated to the fusing point.

665. Cathode Rays carry Negative Charges. — The theory that cathode rays consist of negatively charged particles shot out from the cathode was originally advanced by Crookes. This

view was strongly supported by the fact of magnetic deflection and by the mutual repulsion exerted on each other by two parallel cathodic streams. It was finally shown by experiment that along the path of the cathode rays there is a transport of negative electricity. The experiment was originally devised by Perrin and later modified by Sir J. J. Thomson. The form of the tube used is shown in section in Figure 436. The cathode rays from C pass into the large bulb and fall upon the glass at a point B. In a side branch is placed a metal tube E, with a slit at the inner end, and connected to the earth, to shield the inner metal tube D from any stray electric effects. This latter tube is connected to an electrometer.

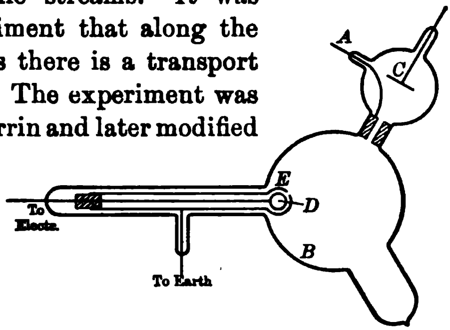


Fig. 436

When the tube is excited the cathode rays travel directly across the large bulb and the electrometer is only slightly affected; but when the rays are deflected by a magnet, so that they enter the slit in the tube E, the electrometer indicates that the tube D has acquired a negative charge. If the rays are deflected further so as to miss the slit, the tube D at once ceases to receive a charge. Cathode rays are therefore at least accompanied by a negative charge of electricity.

In another experiment cathode rays are allowed to pass between two parallel plates E and D (Fig. 437) in a highly exhausted cathode ray tube. When a large difference of potential is produced between the parallel plates by connecting them with the terminals of a battery of a great many storage cells, the beam of rays is deflected from a to b, and the deflection is in the same direction as a negatively charged particle would move in the electric field between

the plates. The facts that cathode rays carry a negative charge and that they are deflected both by a magnetic and an electric field point to the conclusion that they consist of very small negatively charged particles projected from the cathode in straight lines and with a high velocity.

666. Mass of the Ion in Cathode Rays. — By measuring the curvature of the cathode rays when deflected by a magnetic field (Fig. 437), and by balancing the deflection due to a

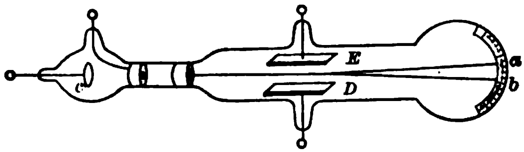


Fig. 437

known magnetic field against that due to a known electrostatic force, it has been found possible to calculate the velocity of the negatively charged particles, and the ratio between the charge e of each particle and its mass m. In this way it has been determined that the average velocity of such particles in a high vacuum is about 3×10^9 cm. per second, or one tenth the velocity of light. The average value of e/m is about 1.8×10^7.

It is of great interest to compare the value of e/m in the case of cathode ions with the value of the same ratio in electrolysis. The greatest value of e/m in electrolysis is in the case of the hydrogen ion, where it is of the order 10^4 (§ 549). Now it has been demonstrated that the charge carried by the negative ion in a high vacuum is the same as the charge carried by the hydrogen ion in electrolysis. But e/m for the negative ion is 1800 times e/m for the hydrogen ion in electrolysis; consequently the mass of the negatively charged cathode ion is only $\frac{1}{1800}$ as great as the mass of the hydrogen ion. For this reason the carrier of the negative charge in cathode rays is called a negative *corpuscle* or an *electron*.

667. Lenard Rays. — Lenard found that it is possible to obtain cathode rays outside an exhausted tube. He used a tube with a metal end opposite the cathode, and a small hole through this was covered with very thin aluminum foil. When cathode rays were directed on this small window, rays issued from it on the outside and produced phosphorescence the same as do cathode rays within the tube. These rays do not proceed far in air at ordinary pressure because they are absorbed. If, however, the window is covered with another exhausted tube, the rays from it form a distinct pencil in the gas at low pressure and carry a negative charge like ordinary cathode rays.

It has been found that a thin piece of metal inside a Crookes tube, when bombarded with cathode rays, has the property of giving off cathode rays from the opposite side, as if the shock of the bombardment were capable of expelling electrons from the opposite surface, when the plate is very thin. Thus the Lenard rays may not be due to electrons from within that have penetrated the aluminum foil, but to a new set driven off from the outside surface because of the impact of charged corpuscles on the inner surface.

668. Roentgen Rays. — In 1895 Roentgen discovered that the impact of cathode rays on glass, and particularly on platinum, gives rise to a new kind of emanations or rays, which affect a sensitized photographic plate like light. They are known as Roentgen rays or X-rays. They differ from cathode rays because they are not deflected by a magnet, and they are able to pass through opaque bodies, such as soda glass, wood, hard rubber, thin sheets of aluminum, and flesh. The more dense metals, like lead, are nearly opaque to them. They travel with the velocity of light, but, unlike light, they do not appear to be refracted in passing from one medium to another.

Fig. 438

The focus tube (Fig. 438) is an effective form for producing Roentgen rays. The cathode is at K, and at the center of curvature of the

cathode is placed obliquely a piece of platinum *A*. The cathode rays are focused on it, and the Roentgen rays have their origin at the point where the cathode stream strikes the platinum and heats it red hot. This platinum plate may also be used as the anode.

669. X-Ray Pictures. — The penetrating power of Roentgen rays depends largely on the pressure within the focus tube. With high exhaustion, and therefore large potential difference between the electrodes of the tube, the rays have high penetrating power. They are then known as "hard rays," and can readily penetrate several centimeters of wood, and even a few millimeters of lead. With a somewhat lower exhaustion, or higher pressure, and smaller potential difference between electrodes, the rays are less penetrating, and are then known as "soft rays."

Fig. 439

Generally speaking the denser substances have the greater absorbing power for X-rays. The possibility of X-ray photographs arises from the variation in this absorbing power. Thus, the bones of the body absorb Roentgen rays more than the flesh; hence fewer rays traverse them. Since Roentgen rays cannot be focused, all photographs taken by means of them are only shadow pictures. As in the case of light, the shadow is the sharper the smaller the area of the source of the radiation.

Figure 439 shows a Roentgen photograph of a gloved hand, in which the ring on the little finger, the two glove-buttons, and the cuff-stud are conspicuous. The flesh is scarcely visible on account of the high penetrating power of the rays used. The photographic plate for the purpose is enclosed in an ordinary plate holder, and the hand is laid on the holder next to the sensitized side.

670. The Fluorescope. — Soon after the discovery of Roentgen rays it was found that certain fluorescent substances, such as platinobarium cyanide and calcium tungstate, become luminous under the action of Roentgen rays. This fact has been turned to account in the construction of the *fluorescope* (Fig. 440), by means of which shadow pictures of concealed objects become visible. An opaque screen is covered on one side with the fluorescent substance; this screen fits into the larger end of a box blackened inside, and having at the other end an opening adapted to fit closely around the eyes, so as to exclude all outside light. When an object, such as the hand, is held against the fluorescent screen and the fluorescope is turned toward the Roentgen tube, the bones are plainly visible on the fluorescent surface inside as darker objects than the flesh because they are more opaque to X-rays. The beating heart may be made visible in a similar manner.

Fig 440

II. IONIZATION OF GASES

671. The Conducting State of a Gas. — The last few sections have been devoted to the phenomena of conductivity in gases without reference to the methods by which a gas may be put into a conducting state. Within the last few years a number of methods have been discovered of increasing the conductivity of gases by artificial means.

It has been found that a gas made artificially conducting may be caused to lose this capacity by bubbling through water, by being forced through a long metal tube of fine bore, or through a plug of glass wool. Further, a gas under the influence of a strong electrostatic field loses its conducting power. The loss of conducting power by filtration and by the action of an electric field shows that the conductivity of a gas is due to the presence of some kind of electrically charged particles mixed with the gas. Also, since a gas in the conducting state exhibits as a whole no

electric charge, the active particles must be charged, some positively and some negatively. In other words, the gas has been *ionized*. But the ions producing conductivity in gases are not the same as the ions in electrolytic conductivity, for the negative cathode ion is only about $\frac{1}{1000}$ of the hydrogen ion in electrolysis.

672. Ionization by Ultra-Violet Light. — A concentrated beam of ultra-violet light may be obtained by converging the light from a carbon electric arc by means of a quartz lens. When such a beam is focused on the negative terminal of a charged Leyden jar, which is separated from the positive just far enough to prevent an automatic discharge, the jar will be discharged. The jar may be charged by an induction coil or an influence machine, and must be equipped with a discharge circuit. If the discharge balls are zinc, the action of the ultra-violet light is the more effective.

When ultra-violet light falls on the surface of an insulated plate of clean zinc, the plate acquires a positive charge; if the plate is charged negatively, it loses its charge; if it is charged positively, no loss of charge occurs. Photo-electric effects may be produced also by the action of ultra-violet light on amalgams of sodium and potassium. They have been shown to be due to the liberation of negative electrons from the metal.

The formation of negative electrons by ultra-violet light renders the air conducting to the extent that a discharge takes place when the air is on the point of breaking down. Nipher has shown that a discharge between terminals does not take place in air at atmospheric pressure until ionization extends across from one terminal to the other.

673. Ionization by Roentgen Rays. — Roentgen rays possess the property of ionizing gases. A well-insulated gold leaf electroscope (§ 483) in dry air will retain its charge for many hours. If, however, a beam of Roentgen rays be caused to stream across the diverging leaf, or the ball at the top, the

leaf will lose its charge and quickly collapse. The air becomes conducting, and the charge, whether positive or negative, leaks away.

The electroscope may also be discharged by leading into it a stream of ionized air. A convenient arrangement for the purpose is shown in Figure 441. Roentgen rays from the tube

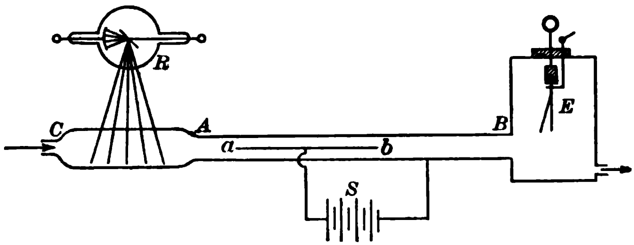


Fig. 441

R fall on the air in the tube *CA*, and the air may be drawn into the electroscope *E* through the metal tube *AB*. The electroscope is not discharged when the Roentgen tube is working unless the air is drawn through it in the direction of the arrows; but as soon as the air is drawn through the charge is lost. The ions produced by the Roentgen rays are therefore carried along with the air. They may be filtered out by means of a plug of glass wool or of cotton; and if an insulated wire *ab* be supported centrally in the metal tube, and a large battery *S* be employed to maintain a potential difference between the wire and the tube, the air will lose its conductivity in passing through this electric field. The ions disappear also by diffusion, for the conductivity of the air persists for a short time only after ionization by Roentgen rays.

674. The Saturation Current. — The rate at which electricity is transported across from one plate to another through ionized gas between them may be studied by means of an arrangement shown in Figure 442. *A* and *C* are two insulated metal plates a few centimeters apart; *A* is connected to one

electrode of a battery *B*, and *C* to one pair of quadrants of an electrometer *E*. The other pair of quadrants is connected to earth; also the other electrode of the battery.

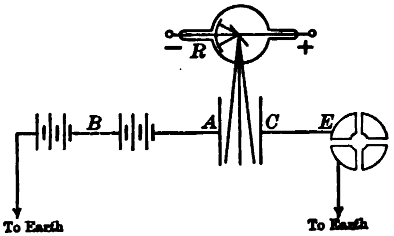


Fig 442

When the tube *R* is excited, a beam of Roentgen rays passes between the two plates, and *E* indicates at once that it is receiving a charge. If *A* is connected to the positive electrode *B*, then *C* receives a positive charge. The rays thus cause a transport of electricity across from *A* to *C*, and the sign of the charge communicated to *C* is the same as the sign of *A*.

The rate at which the electrometer deflection increases is taken as a measure of the current through the gas between the plates. By measuring the current for different potential differences between the plates, the relation between the two quantities may be obtained by plotting them as coördinates. For small values of the E. M. F. of the battery *B*, the curve (Fig. 443) is a straight line, or the current is proportional to the E. M. F. and therefore follows Ohm's law. As the E. M. F. of the battery increases, the current does not increase as fast as Ohm's law requires; and finally a condition is reached where there is no further increase of current for a considerable increase of the E. M. F. The curve then becomes flat, and the corresponding current is called the *saturation current* for the gas under the given conditions.

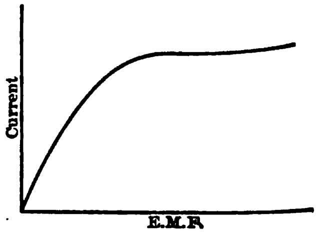


Fig. 443

The current through an ionized gas differs in another respect from the current through an electrolyte. When the

distance between two electrodes immersed in mercury or in an electrolyte is increased, the resistance increases and the current decreases; but when the distance between the plates in an ionized gas is increased, the saturation current also is increased.

675. Theory of Ionization.—The theory of ionization supposes that when a gas is ionized the molecules are broken up into positively and negatively charged carriers of electricity or ions. The process is called ionization. The two ions resulting from the breaking in two of the molecule have equal charges of opposite signs, as in electrolytic dissociation. The transport of electricity through gases is effected by the movement of ions under the influence of an electric field.

Gaseous ions may recombine by mutual attraction when they come within the sphere of each other's influence. Their electrical properties are then neutralized and they disappear as ions.

The greater the potential difference between two plates where ionization takes place, the faster the charged carriers move, the positive toward lower potential and the negative toward higher, and the number of ions recombining is reduced in due proportion. The current is proportional to the number of ions reaching the plates per second. When the voltage reaches a certain value, no recombination takes place because the ions are hurried on so fast toward the two plates. The ions are then transferred to the plates as fast as they are formed, and a higher voltage cannot increase the rate of transfer. Hence the saturation current.

Observations on the diffusion of ions have led to the conclusion that both positive and negative ions at atmospheric pressure are the nuclei of a cluster of molecules. The ionization consists in separating a negative electron from a neutral molecule, the positive ion consisting of the remainder of the molecule with its positive charge. At low pressures the negative ion is the same as the electron, but at higher

pressures and in moist gases both ions gather about them molecules of gas and water. Hence at low pressures and in dry gases the negative ion diffuses much more rapidly than the positive, while at high pressures, and especially in the presence of moisture, their rates of diffusion are more nearly the same.

Similar explanations apply to the fact that the velocity of negative ions for the same potential gradient is greater than that of positive ions; also to the fact that the velocities of both are greater in light gases like hydrogen than in heavy ones.

676. Condensation of Water Vapor by Ions. — It was supposed for some time that water vapor could not be made to condense by moderate expansion in a dustless atmosphere. Dust particles serve as nuclei for the condensation of moisture. But it was shown by Wilson that when a mass of saturated gas, free from dust but containing ions, is caused to expand 25 per cent of its original volume, a cloud is produced. The gas becomes supersaturated on account of the fall of temperature due to the expansion; in the absence of both dust and ions, the supersaturation is not sufficient to produce drops of liquid; but when ions are present, they serve as nuclei for the condensation of water.

Sir J. J. Thomson and others have made use of this fact to determine the number of ions present and the elementary charge on each. If the ions are not numerous, each one acts as a nucleus for a single drop, and the number of drops formed is the same as the number of ions present.

Sir George Stokes obtained an expression for the velocity with which a drop falls in terms of the diameter of the drop, the acceleration of gravity, and the viscosity of the gas. Hence the observed rate of subsidence of the cloud gives the mean diameter of the drops. The mass of water condensed from each cubic centimeter can be calculated from the cooling corresponding to the given adiabatic expansion

and the heat of evaporation of water. Then knowing the mass of water condensed per cubic centimeter, and the mean diameter of the drops, the number of drops, which is the same as the number of ions, or a simple sub-multiple of them, can be readily calculated.

If, further, all the ions present be withdrawn by an electric field, the total charge can be measured. In this way one obtains the charge on each ion. Sir J. J. Thomson obtained the value

$$e = 3.4 \times 10^{-10} \text{ electrostatic units.}$$

The latest and most consistent value is that of Millikan,

$$e = 4.774 \times 10^{-10} \text{ electrostatic units.}$$

This elementary charge appears to be independent of the kind of gas and of other conditions.

III. RADIOACTIVITY

677. Discovery of Radioactivity. — In 1896 Becquerel discovered that the double sulphate of potassium and uranium emits a radiation which has an effect on a photographic plate similar to that of Roentgen rays. It was also found that these radiations cause the discharge of electrified bodies, and Rutherford demonstrated that they produce this effect by ionizing a gas, similar to the ionization produced by Roentgen rays.

In the arrangement of Figure 442 the focus tube may be removed and a compound of uranium may be sprinkled on one of the parallel plates arranged in a horizontal position; the electroscope will be charged the same as if the air had been ionized by the passage of X-rays. Uranium and all other bodies possessing similar properties are said to be *radioactive*. Radioactivity is characteristic of all compounds of uranium, the intensity of the radiation depending on the amount of uranium present.

The examination of other bodies for radioactive properties has shown that the element thorium and its compounds are also radioactive, though distinctly less so than uranium and its compounds, especially in their photographic effect.

678. The Discovery of Radium. — The investigations in radioactivity revealed the fact that certain samples of pitchblende were more highly radioactive than either thorium or uranium. This observation suggested the presence in pitchblende of some substance of far greater radioactivity than uranium itself. M. and Mme. Curie succeeded by purely chemical means in separating from pitchblende a new element, which they named polonium, and subsequently another which was named *radium*. The latter in a pure state is a million times more active than uranium. Up to the present time it is the most remarkable radioactive substance known. The study of its properties has led to a knowledge that hitherto unknown and unsuspected processes are going on in nature on a wide scale.

Radium occurs in a number of minerals and is widely distributed, but its chief source is the pitchblende found in Bohemia. It forms only an almost infinitesimal part of the pitchblende, a ton of this material yielding only a few milligrams of pure radium. The radium commonly used for investigation is not the element, but the bromide of radium. It forms other salts also, such as the chloride and the sulphate, and all of them are radioactive.

679. Fluorescence excited by Radium. — Salts of radium have the remarkable property of exciting strong fluorescence. When a glass tube containing bromide of radium is held against the bottom of the fluoroscope, an area of light, with poorly defined outlines, can be seen on the screen within. The radium bromide gives off emanations which produce fluorescence when they strike the screen. A diamond exposed to highly active radium bromide glows with a soft light in the dark.

Sir William Crookes has invented an instrument, called a *spinthariscope*, which beautifully illustrates fluorescence by radium. In one end of a short brass tube is fixed a small screen covered with fluorescent zinc sulphide, and near it is supported a bit of radium bromide. In the other end of the tube is a simple magnifier for viewing the screen. When the eye is sufficiently sensitive, not only is the screen seen to be luminous, but the whole area near the radium presents a most interesting appearance of a boiling, scintillating motion. Wherever the emanations from the radium strike the screen, a slight flash of light is produced. The field of view appears therefore to be violently agitated, and at the same time it emits light.

680. Heat produced by Radium. — An altogether new and remarkable property, exhibited by the salts of radium, was discovered by Curie and Laborde; they are always maintained at a temperature several degrees above that of the surrounding air. By reason of this excess of temperature they are incessantly radiating heat. It has been calculated that a gram of pure radium would emit heat at the rate of 100 calories per hour. Thorium also emits heat, though in a minor degree.

The explanation of this generation of heat is the readiness with which the radiations, starting within the radioactive body, are absorbed by itself with the transformation of their kinetic energy into heat.

681. Three Kinds of Rays. — The early discovery that the radiations from uranium are of a complex nature was made by Rutherford. Three kinds of radiations from radioactive bodies have since been differentiated from one another. They are known as the α, the β, and the γ rays. In general those which are the most penetrating and have the greatest photographic effect are the least efficient in producing ionization.

The α rays are readily absorbed by metals and even by gases; that is, they have little penetrating power. The most penetrating α rays are absorbed by a thickness of 0.01 cm. of aluminum, or by a sheet of ordinary writing paper. They are, however, the most effective ionizers, though the ionization

produced by them is limited to a few centimeters from their source, because they are readily absorbed by the air at atmospheric pressure. The α rays consist of positively charged particles emitted with a velocity about a tenth of the speed of light, very likely helium atoms.

The β rays are much more active on a photographic plate than the α rays. Most of the photographic action of the rays from radium is due to β rays. It has been found that the β rays are negatively charged. When they fall on an insulated plate connected with an electrometer, the latter acquires a negative charge.

The γ rays have very small photographic activity; in fact, those from uranium and thorium produce no observable photographic effect. The γ rays given out by uranium, thorium, and radium are extremely penetrating. They are able to ionize a gas and make it conducting after passing through a thickness of 30 cm. of iron. They resemble Roentgen rays from a "hard" X-ray tube, but are much more penetrating. Though they always occur in conjunction with β rays, they do not appear to carry an electric charge.

682. Magnetic Deviation of Three Kinds of Rays. — The three types of rays exhibit very different behavior with respect to deflection by a magnetic field. This difference in behavior is illustrated in Figure 444, which is due to Mme. Curie. The radium is placed at the bottom of a small hole in a block of lead L. The hole is so narrow that only a thin pencil of the three types of rays escapes in a vertical direction. A strong magnetic field is applied in a direction perpendicular to the plane of the paper, and so that the lines of force run away from the observer. The β rays are strongly deflected to the right, the α rays very slightly to the left, while the γ rays are not affected in the least. The

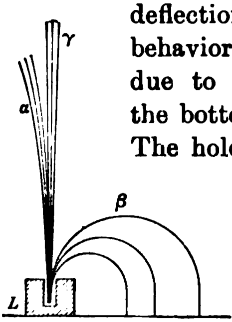


Fig. 444

deviation of the α rays is greatly exaggerated in the figure in order to show it at all. The fact that the α and β rays suffer deviations in opposite directions shows not only that they are charged particles, but that they are oppositely charged, the former positively and the latter negatively.

683. Emanation from Radioactive Bodies. — In an examination of the irregularity of the radiation from thorium Rutherford discovered that thorium is incessantly giving off a material emanation which has the properties of a radioactive gas. Radium also gives off an emanation with properties similar to those of the thorium emanation, but uranium and polonium are not known to give off any emanation.

This emanation diffuses through gas and porous bodies, and can be condensed in a tube immersed in liquid air at about $-150°$. It can be carried away by a current of air, and is capable of ionizing a gas and acting on a photographic plate. It may be bubbled through water or drawn through cotton without losing its ionizing power.

The activity of the emanation dies away rapidly, that from thorium losing half its activity in about a minute; the corresponding period for the radium emanation is about 3.7 days.

When the emanation from radium or thorium falls on a solid body, its surface becomes coated with a very thin deposit of extremely radioactive matter. It is invisible, but can be dissolved off by certain acids; when the solvent is evaporated, the active matter is left as a residue. It emits radiations affecting a photographic plate and ionizes a gas. This "excited" activity gradually dies away in a manner similar to that for the decay of the activity of the emanation itself.

684. The Cause of Radioactive Changes. — To account for the phenomena of radioactivity the theory is held that radioactive substances are gradually and spontaneously undergoing atomic changes or disintegration.

In 1900 Crookes separated from uranium by a simple

chemical process a constituent far more active than uranium itself, and the separation left the uranium photographically inactive. The new constituent Crookes named Uranium X. Becquerel discovered subsequently that in the course of time the uranium recovers its usual activity while the uranium X loses it entirely. Later Rutherford and Soddy separated a similar active constituent from thorium and called it Thorium X. Thorium also recovers its activity and thorium X loses it.

Further, Ramsay and Soddy collected radium emanation in a vacuum tube and examined its spectrum. It was new and was considered to be the spectrum of the emanation itself. After a few days, however, the spectrum of helium made its appearance, indicating that the final product of the atomic changes in the case of radium is helium.

According to the disintegration theory, the atom of radium, as a typical example, is a complex structure consisting of systems of positively and negatively charged particles in rapid rotation and in somewhat unstable equilibrium. On account of this instability one of the α particles is suddenly expelled with great velocity. The structure remaining constitutes an atom of the emanation. This is also unstable and expels another α particle, and the process continues through the successive changes.

It is quite possible that the energy which radioactive bodies are continually giving off owes its origin to some such internal atomic change as the theory supposes. While the atom of radium is radiating energy in the form of heat and active rays, it is at the same time losing potential energy. The hypothesis suggests that all matter may be in process of similar changes of disintegration, and that these changes have not been observed because the expelled particles, except in the case of radioactive bodies, are not expelled with sufficient energy to ionize a gas or to produce photographic action.

INDEX

Numbers refer to pages